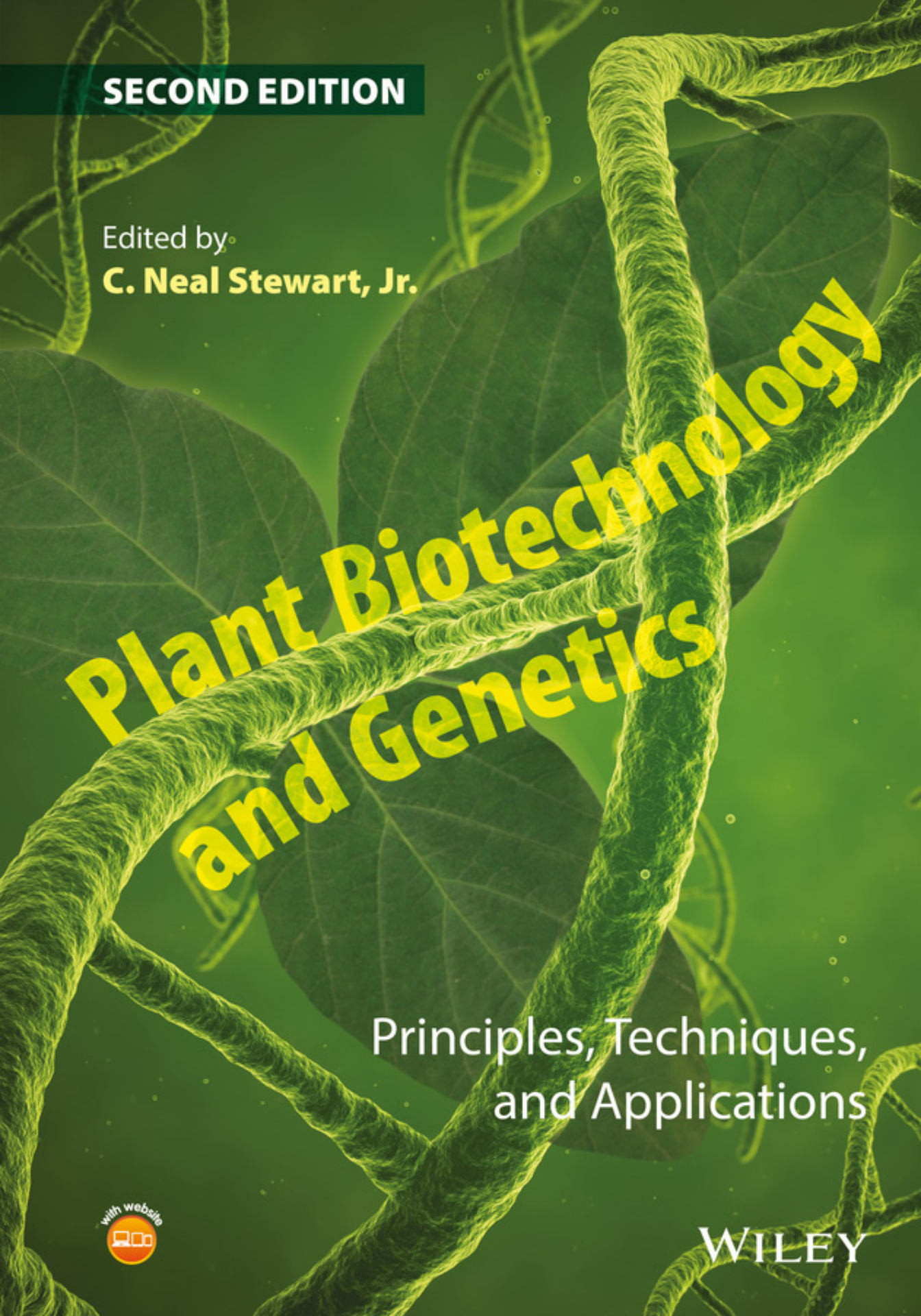

SECOND EDITION
Edited by
C. Neal Stewart, Jr.
Plant Biotechnology and Genetics
Principles, Techniques, and Applications
with website
WILEY

PLANT BIOTECHNOLOGY AND GENETICS

PLANT BIOTECHNOLOGY AND GENETICS

Principles, Techniques, and Applications

Second Edition

Edited by
C. Neal Stewart, Jr.

WILEY

Copyright © 2016 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Stewart, C. Neal, Jr. author.
Plant biotechnology and genetics : principles, techniques, and applications / edited by C. Neal Stewart, Jr. – Second edition.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-118-82012-4 (hardback)
I. Title.
[DNLM: 1. Biotechnology–methods. 2. Plants, Genetically Modified–genetics. 3. Genetic Enhancement–methods. TP 248.27.P55]
TP248.27.P55
660.6′5–dc23

2015033118

Cover image courtesy of Jennifer Hinds

Set in 10/12pt Times by SPi Global, Pondicherry, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

2 2016

To the next generation of pioneers.

CONTENTS

Foreword xvi

Contributors xviii

Preface xx

1. The Impact of Biotechnology on Plant Agriculture 1
Graham Brookes

1.0 Chapter Summary and Objectives 1
1.0.1 Summary 1
1.0.2 Discussion Questions 1
1.1 Introduction 1
1.2 Cultivation of Biotechnology (GM) Crops 2
1.3 Why Farmers Use Biotech Crops 4
1.4 GM's Effects on Crop Production and Farming 7
1.5 How the Adoption of Plant Biotechnology has Impacted the Environment 8
1.5.1 Environmental Impacts from Changes in Insecticide and Herbicide Use 8
1.5.2 Impact on GHG Emissions 11
1.6 Conclusions 13
Life Box 1.1 Norman E. Borlaug 14
Life Box 1.2 Mary-Dell Chilton 15
Life Box 1.3 Robert T. Fraley 17
References 19

2. Mendelian Genetics and Plant Reproduction 20
Matthew D. Halfhill and Suzanne I. Warwick

2.0 Chapter Summary and Objectives 20
2.0.1 Summary 20
2.0.2 Discussion Questions 20
2.1 Overview of Genetics 20
2.2 Mendelian Genetics 23
2.2.1 Law of Segregation 26
2.2.2 Law of Independent Assortment 26
2.3 Mitosis and Meiosis 27
2.3.1 Mitosis 29
2.3.2 Meiosis 29
2.3.3 Recombination 30
2.3.4 Cytogenetic Analysis 31
2.3.5 Mendelian Genetics and Biotechnology Summary 32
2.4 Plant Reproductive Biology 32
2.4.1 History of Research in Plant Reproduction 32
2.4.2 Mating Systems 32
2.4.3 Hybridization and Polyploidy 36
2.4.4 Mating Systems and Biotechnology Summary 38
2.5 Conclusion 38

Life Box 2.1 Richard A. Dixon 39
Life Box 2.2 Michael L. Arnold 40
References 42

3. Plant Breeding 43
Nicholas A. Tinker and Elroy R. Cober

3.0 Chapter Summary and Objectives 43
3.0.1 Summary 43
3.0.2 Discussion Questions 43
3.1 Introduction 44
3.2 Central Concepts in Plant Breeding 45
3.2.1 Simple vs. Complex Inheritance 45
3.2.2 Phenotype vs. Genotype 46
3.2.3 Mating Systems, Varieties, Landraces, and Pure Lines 47
3.2.4 Other Topics in Population and Quantitative Genetics 49
3.2.5 The Value of a Plant Variety Depends on Many Traits 51
3.2.6 A Plant Variety Must Be Environmentally Adapted 51
3.2.7 Plant Breeding is a Numbers Game 52
3.2.8 Plant Breeding is an Iterative and Collaborative Process 52
3.2.9 Diversity, Adaptation, and Ideotypes 53
3.2.10 Other Considerations 56
3.3 Objectives in Plant Breeding 56
3.4 Methods of Plant Breeding 57
3.4.1 Methods of Hybridization 58
3.4.2 Self-Pollinated Species 58
3.4.3 Outcrossing Species 63
3.4.4 Clonally Propagated Species 67
3.5 Breeding Enhancements 68
3.5.1 Doubled Haploidy 68
3.5.2 Marker-Assisted Selection 68
3.5.3 Mutation Breeding 70
3.5.4 Apomixis 71
3.6 Conclusions 71
Life Box 3.1 Gurdev Singh Khush 72
Life Box 3.2 P. Stephen Baenziger 74
Life Box 3.3 Steven D. Tanksley 75
References 77

4. Plant Development and Physiology 78
Glenda E. Gillaspy

4.0 Chapter Summary and Objectives 78
4.0.1 Summary 78
4.0.2 Discussion Questions 78
4.1 Plant Anatomy and Morphology 79
4.2 Embryogenesis and Seed Germination 80
4.2.1 Gametogenesis 80
4.2.2 Fertilization 82
4.2.3 Fruit Development 83
4.2.4 Embryogenesis 83
4.2.5 Seed Germination 85
4.2.6 Photomorphogenesis 85
4.3 Meristems 86
4.3.1 Shoot Apical Meristem 86

4.3.2 Root Apical Meristem and Root Development 88
4.4 Leaf Development 89
4.4.1 Leaf Structure 89
4.4.2 Leaf Development Patterns 91
4.5 Flower Development 92
4.5.1 Floral Evocation 92
4.5.2 Floral Organ Identity and the ABC Model 93
4.6 Hormone Physiology and Signal Transduction 94
4.6.1 Seven Plant Hormones and Their Actions 94
4.6.2 Plant Hormone Signal Transduction 96
4.7 Conclusions 100
Life Box 4.1 Deborah Delmer 100
Life Box 4.2 Natasha Raikhel 102
Life Box 4.3 Brenda S.J. Winkel 103
References 105

5. Tissue Culture: The Manipulation of Plant Development 107
Vinitha Cardoza

5.0 Chapter Summary and Objectives 107
5.0.1 Summary 107
5.0.2 Discussion Questions 107
5.1 Introduction 107
5.2 History of Tissue Culture 108
5.3 Media and Culture Conditions 109
5.3.1 Basal Media 109
5.3.2 Growth Regulators 110
5.4 Sterile Technique 111
5.4.1 Clean Equipment 111
5.4.2 Surface Sterilization of Explants 112
5.5 Culture Conditions and Vessels 113
5.6 Culture Types and Their Uses 113
5.6.1 Callus and Somatic Embryo Culture 113
5.6.2 Cell Suspension Cultures 117
5.6.3 Anther/Microspore Culture 119
5.6.4 Protoplast Culture 119
5.6.5 Somatic Hybridization 120
5.6.6 Embryo Culture 120
5.6.7 Meristem Culture 121
5.7 Regeneration Methods of Plants in Culture 121
5.7.1 Organogenesis 121
5.7.2 Somatic Embryogenesis 123
5.7.3 Synthetic Seeds 123
5.8 Rooting of Shoots 123
5.9 Acclimation 124
5.10 Problems that can Occur in Tissue Culture 124
5.10.1 Culture Contamination 124
5.10.2 Hyperhydricity 124
5.10.3 Browning of Explants 124
5.11 Conclusions 125
Acknowledgments 125
Life Box 5.1 Glenn Burton Collins 125
Life Box 5.2 Martha S. Wright 127
Life Box 5.3 Vinitha Cardoza 128
References 129

6. Molecular Genetics of Gene Expression **133**
Maria Gallo and Alison K. Flynn

6.0 Chapter Summary and Objectives 133
6.0.1 Summary 133
6.0.2 Discussion Questions 133
6.1 The Gene 133
6.1.1 DNA Coding for a Protein via the Gene 133
6.1.2 DNA as a Polynucleotide 134
6.2 DNA Packaging into Eukaryotic Chromosomes 134
6.3 Transcription 135
6.3.1 Transcription of DNA to Produce Messenger Ribonucleic Acid 135
6.3.2 Transcription Factors 140
6.3.3 Coordinated Regulation of Gene Expression 140
6.3.4 Chromatin as an Important Regulator of Transcription 141
6.3.5 Regulation of Gene Expression by DNA Methylation 142
6.3.6 RNA-Directed Gene Silencing by Small RNAs 143
6.3.7 Processing to Produce Mature mRNA 143
6.4 Translation 144
6.4.1 Initiation of Translation 147
6.4.2 Elongation Phase of Translation 147
6.4.3 Translation Termination 147
6.5 Protein Postranslational Modification 147
Life Box 6.1 Maarten Chrispeels 150
Life Box 6.2 David W. Ow 152
References 154

7. Plant Systems Biology **155**
Wusheng Liu and C. Neal Stewart, Jr.

7.0 Chapter Summary and Objectives 155
7.0.1 Summary 155
7.0.2 Discussion Questions 155
7.1 Introduction 155
7.2 Defining Plant Systems Biology 157
7.3 Properties of Plant Systems 158
7.4 A Framework of Plant Systems Biology 159
7.4.1 Comprehensive Quantitative Data Sets 160
7.4.2 Network Analysis 161
7.4.3 Dynamic Modeling 161
7.4.4 Exploring Systems and Models Toward Refinement 161
7.5 Disciplines and Enabling Tools of Plant Systems Biology 162
7.5.1 Plant Genomics 162
7.5.2 Plant Transcriptomics 166
7.5.3 Plant Proteomics 168
7.5.4 Plant Metabolomics 170
7.5.5 Bioinformatics 172
7.6 Conclusions 176
Life Box 7.1 C. Robin Buell 177
Life Box 7.2 Zhenbiao Yang 178
References 179

8. Recombinant DNA, Vector Design, and Construction **181**
Mark D. Curtis and David G.J. Mann

8.0 Chapter Summary and Objectives 181
8.0.1 Summary 181
8.0.2 Discussion Questions 181

8.1 DNA Modification 181
8.2 DNA Vectors 186
8.2.1 DNA Vectors for Plant Transformation 188
8.2.2 Components for Efficient Gene Expression in Plants 190
8.3 Greater Demands Lead to Innovation 192
8.3.1 "Modern" Cloning Strategies 192
8.4 Vector Design 197
8.4.1 Vectors for High-Throughput Functional Analysis 197
8.4.2 Vectors for Gene Down-Regulation Using RNA Interference (RNAi) 199
8.4.3 Expression Vectors 199
8.4.4 Vectors for Promoter Analysis 200
8.4.5 Vectors Derived from Plant Sequences 201
8.4.6 Vectors for Multigenic Traits 203
8.5 Targeted Transgene Insertions 204
8.6 Prospects 205
Life Box 8.1 Wayne Parrott 206
Life Box 8.2 David Mann 207
References 208

9. Genes and Traits of Interest 211
Kenneth L. Korth

9.0 Chapter Summary and Objectives 211
9.0.1 Summary 211
9.0.2 Discussion Questions 211
9.1 Introduction 212
9.2 Identifying Genes of Interest via Genomics and other Omics Technologies 212
9.3 Traits for Improved Crop Production Using Transgenics 214
9.3.1 Herbicide Resistance 215
9.3.2 Insect Resistance 218
9.3.3 Pathogen Resistance 220
9.3.4 Traits for Improved Products and Food Quality 222
9.4 Conclusion 227
Life Box 9.1 Dennis Gonsalves 227
Life Box 9.2 Ingo Potrykus 229
References 231

10. Promoters and Marker Genes 233
Wusheng Liu, Brian Miki and C. Neal Stewart, Jr.

10.0 Chapter Summary and Objectives 233
10.0.1 Summary 233
10.0.2 Discussion Questions 233
10.1 Introduction 234
10.2 Promoters 234
10.2.1 Constitutive Promoters 235
10.2.2 Tissue-Specific Promoters 236
10.2.3 Inducible Promoters 237
10.2.4 Synthetic Promoters 239
10.3 Marker Genes 239
10.3.1 Selectable Marker Genes 242
10.3.2 Reporter Genes 246
10.4 Marker-Free Strategies 250
10.5 Conclusions 254
Life Box 10.1 Fredy Altpeter 255
Life Box 10.2 Taniya Dhillon 257
References 259

11. Transgenic Plant Production **262**
John J. Finer

11.0 Chapter Summary and Objectives 262
11.0.1 Summary 262
11.0.2 Discussion Questions 262
11.1 Overview of Plant Transformation 263
11.1.1 Introduction 263
11.1.2 Basic Components for Successful Gene Transfer to Plant Cells 263
11.2 *Agrobacterium Tumefaciens* 265
11.2.1 History of *Agrobacterium* Research 266
11.2.2 Use of the T-DNA Transfer Process for Transformation 268
11.2.3 Optimizing Delivery and Broadening the Taxonomical Range of Targets 269
11.2.4 Strain and Cultivar Compatibility 270
11.2.5 Agroinfiltration 271
11.2.6 *Arabidopsis* Floral Dip (Clough and Bent 1998) 271
11.3 Particle Bombardment 272
11.3.1 History of Particle Bombardment 272
11.3.2 The Fate of the Introduced DNA into Plant Cells 274
11.3.3 The Power and Problems of Direct DNA Introduction 275
11.3.4 Improvements in the Control of Transgene Expression 276
11.4 Other Methods of Transformation 276
11.4.1 The Need for Additional Technologies 276
11.4.2 Protoplasts 277
11.4.3 Whole Tissue Electroporation 278
11.4.4 Silicon Carbide Whiskers 278
11.4.5 Viral Vectors 278
11.4.6 Laser Micropuncture 279
11.4.7 Nanofiber Arrays 279
11.5 The Rush to Publish 280
11.5.1 Controversial Reports of Plant Transformation 280
11.5.2 Criteria to Consider in Judging Novel Plant Transformation Methods 284
11.6 A Look to the Future 286
Life Box 11.1 Ted Klein 286
Life Box 11.2 John Finer 287
Life Box 11.3 Kan Wang 289
References 291

12. Analysis of Transgenic Plants **293**
C. Neal Stewart, Jr.

12.0 Chapter Summary and Objectives 293
12.0.1 Summary 293
12.0.2 Discussion Questions 293
12.1 Essential Elements of Transgenic Plant Analysis 293
12.2 Assays for Transgenicity, Insert Copy Number, and Segregation 295
12.2.1 Polymerase Chain Reaction 295
12.2.2 Quantitative PCR 295
12.2.3 Southern (DNA) Blot Analysis 296
12.2.4 Segregation Analysis of Progeny 300
12.3 Transgene Expression 301
12.3.1 Transcript Abundance 301
12.3.2 Protein Abundance 302
12.4 Knockdown or Knockout Analysis Rather than Overexpression Analysis 304
12.5 The Relationship Between Molecular Analyses and Phenotype 305

Life Box 12.1 Hong S. Moon 305
Life Box 12.2 Neal Stewart 306
Life Box 12.3 Nancy A. Reichert 308
References 310

13. Regulations and Biosafety 311
Alan McHughen

13.0 Chapter Summary and Objectives 311
13.0.1 Summary 311
13.0.2 Discussion Questions 311
13.1 Introduction 311
13.2 History of Genetic Engineering and Its Regulation 313
13.3 Regulation of GM Plants 315
13.3.1 New Technologies 316
13.3.2 US Regulatory Agencies and Regulations 317
13.3.3 European Union 319
13.3.4 Canada 321
13.3.5 International Perspectives 321
13.4 Regulatory Flaws and Invalid Assumptions 323
13.4.1 Conventional Plant Breeding has Higher Safety than Biotechnology-Derived GM 324
13.4.2 GMOs Should Be Regulated Because They're GMOs and Un-natural 324
13.4.3 Even though Product Risk is Important, It is Reasonable that Process (GMO) Should Trigger Regulation 324
13.4.4 Since GM Technology is New, It Might Be Hazardous and Should Be Regulated 325
13.4.5 If We Have a Valid Scientific Test, Then It Should Be Used in Regulations 326
13.4.6 Better Safe than Sorry: Overregulation is Better than Underregulation 326
13.5 Conclusion 327
Life Box 13.1 Alan McHughen 328
Life Box 13.2 Raymond D. Shillito 329
References 331

14. Field Testing of Transgenic Plants 333
Detlef Bartsch, Achim Gathmann, Christiane Saeglitz and Arti Sinha

14.0 Chapter Summary and Objectives 333
14.0.1 Summary 333
14.0.2 Discussion Questions 333
14.1 Introduction 334
14.2 Environmental Risk Assessment Process 334
14.2.1 Initial Evaluation (Era Step 1) 334
14.2.2 Problem Formulation (ERA Step 2) 335
14.2.3 Controlled Experiments and Gathering of Information (ERA Step 3) 335
14.2.4 Risk Evaluation (ERA Step 4) 335
14.2.5 Progression through a Tiered Risk Assessment 335
14.3 An Example Risk Assessment: The Case of Bt Maize 336
14.3.1 Effect of Bt Maize Pollen on Nontarget Caterpillars 337
14.3.2 Statistical Analysis and Relevance for Predicting Potential Adverse Effects on Butterflies 339
14.4 Proof of Safety Versus Proof of Hazard 340
14.5 Modeling the Risk Effects on a Greater Scale 340
14.6 Proof of Benefits: Agronomic Performance 341
14.7 Conclusions 342
Life Box 14.1 Tony Shelton 343
Life Box 14.2 Detlef Bartsch 344
References 346

15. Intellectual Property in Agricultural Biotechnology: Strategies for Open Access **347**
Monica Alandete-Saez, Cecilia Chi-Ham, Gregory Graff, Sara Boettiger and Alan B. Bennett

15.0 Chapter Summary and Objectives 347
15.0.1 Summary 347
15.0.2 Discussion Questions 347
15.1 Intellectual Property and Agricultural Biotechnology 348
15.1.1 What is Intellectual Property? 349
15.1.2 What is a Patent? 349
15.2 The Relationship Between Intellectual Property and Agricultural Research 351
15.3 Patenting Plant Biotechnology: Has an Anti-Commons Developed? 352
15.3.1 Transformation Methods 352
15.3.2 Selectable Markers 353
15.3.3 Promoters 354
15.3.4 Subcellular Localization 354
15.3.5 The Importance of Combining IP-Protected Components in Transgenic Crops 355
15.4 What is Freedom to Operate (FTO)? 355
15.4.1 The Importance of FTO 355
15.4.2 FTO Case Study: the Tomato E8 Promoter 356
15.5 Strategies for Open Access 358
15.6 Conclusions 359
Life Box 15.1 Alan Bennett 360
Life Box 15.2 Maud Hinchee 361
References 363

16. Why Transgenic Plants Are So Controversial **366**
Jennifer Trumbo and Douglas Powell

16.0 Chapter Summary and Objectives 366
16.0.1 Summary 366
16.0.2 Discussion Questions 366
16.1 Introduction 367
16.1.1 The Frankenstein Backdrop 367
16.1.2 Agricultural Innovations and Questions 367
16.2 Perceptions of Risk 368
16.3 Responses of Fear 370
16.4 Feeding Fear: Case Studies 372
16.4.1 Pusztai's Potatoes 372
16.4.2 Monarch Butterfly Flap 373
16.5 How Many Benefits are Enough? 373
16.6 Continuing Debates 375
16.6.1 Process vs. Product 375
16.6.2 Health Concerns 375
16.6.3 Environmental Concerns 376
16.6.4 Consumer Choice 376
16.7 Business and Control 376
16.8 Conclusions 377
Life Box 16.1 Tony Conner 378
Life Box 16.2 Channapatna S. Prakash 379
References 381

17. The Future: Advanced Plant Biotechnology, Genome Editing, and Synthetic Biology **383**
Wusheng Liu and C. Neal Stewart, Jr.

17.0 Chapter Summary and Objectives 383
17.0.1 Summary 383
17.0.2 Discussion Questions 383

17.1 Introduction: The Birth of Synthetic Biology 384
17.2 Defining Synthetic Biology for Plants 385
17.2.1 Design Cycles of Synthetic Biology 385
17.2.2 Foundations of Synthetic Biology 387
17.2.3 Components of Plant Synthetic Biology 388
17.3 Enabling Tools for Plant Synthetic Biology 389
17.3.1 Computer-Aided Design 389
17.3.2 Synthetic Promoters 389
17.3.3 Precise Genome Editing 389
17.4 Synthetic Biology Applications in Plants 393
17.4.1 Synthetic Inducible Promoters 394
17.4.2 A Device for Monitoring Auxin-Induced Plant IAA Degradation in Yeast 395
17.4.3 Circuits for Phytosensing of Explosives or Bacterial Pathogens in Transgenic Plants 395
17.5 Conclusions 397
Life Box 17.1 Joshua Yuan 397
Life Box 17.2 Wusheng Liu 398
References 399

Index **402**

FOREWORD

An international (but widely unnoticed) race took place in the mid-1970s to understand how *Agrobacterium tumefaciens* caused plant cells to grow rapidly into a gall that produced its favorite substrates—called "opines." Belgian, German, Australian, French, and US groups were at the forefront of different aspects of the puzzle. By 1977, it was clear that gene transfer from the bacterium to its plant host was the secret, and that the genes from the bacterium were functioning to alter characteristics of the plant cells. Participants in the race as well as observers began to speculate that we might exploit the capability of this cunning bacterium in order to get plants to produce our favorite substrates. Small startup companies and multinational corporations took notice and began to work with *Agrobacterium* and other means of gene transfer to plants. One by one the problems were dealt with, and each step in the use of *Agrobacterium* for the genetic engineering of a tobacco plant was demonstrated.

As I look back to those early experiments, I see that we have come a long way since the birth of plant biotechnology, which most of us who served as midwives would date from the Miami Winter Symposium of January 1983. The infant technology was weak and wobbly, but its viability and vitality were already clear. Its growth and development were foreseeable although not predictable in detail. I thought that the difficult part was behind us, and now (as I used to predict at the end of my lectures) the main challenge would be thinking of what genes we might use to bring about desired changes in crop plants. Unseen at that early date were the interesting problems, some technical and some of other kinds, to be encountered and overcome.

To my surprise, one of the biggest challenges turned out to be tobacco, which worked so well that it made us cocky. Tobacco was the guinea pig of the plant kingdom in 1983. This plant has an uncanny ability to reproduce a new plant from (almost) any of its cells. We practiced our gene transfer experiments on tobacco cells with impunity, and we could coax transgenic plants to develop from almost any cell into which *Agrobacterium* had transferred our experimental gene. This ease of regeneration of tobacco did not prepare us for the real world, whose principal food crops (unlike tobacco) were monocots—corn, wheat, rice, sorghum, and millet—to which the technology would ultimately need to be applied. Regeneration of these monocot plants from certain rare cells would be needed, and gene transfer to those very cells must be achieved. This process took years of research, and solutions were unique for each plant. In addition, much of the work was performed in small or large biotech companies, which sought to block competitors by applying for patent protection on methods they developed. Thus, still other methods had to be developed if licensing was not an option.

Another challenge we faced was bringing about expression of the "transgenes" we introduced into the plant cell. We optimistically supposed that any transgene, if given a plant gene promoter, would function in plants. After all, in 1983 the first gene everyone tried, the one coding for neomycin phosphotransferase II, had worked beautifully! The gene encoding a *Bacillus thuringiensis* insecticidal protein (nicknamed Bt, among other things, in the lab) was to teach us humility. Considerable ingenuity was needed to figure out why the Bt gene refused to express properly in the plant, and what to do about it. In the end, we learned to avoid many problems by using an artificial copy of this Bt gene constructed from plant-preferred codons. Although we thought of the genetic code as universal, as a practical matter, correct and fluent gene translation turned out to require, where a choice of codons was provided, that we use the plant's favorites.

An entirely new problem was how to determine product safety. Once the transgenic plant was performing properly, how should it be tested for any unforeseen properties that might conceivably make it harmful, toxic, allergenic, weedy (i.e., a pest in subsequent crops grown in the field), or disagreeable in any other way one could imagine? Ultimately, as they gained experience with these new products, regulatory agencies developed protocols for testing transgenic plants. The transgene must be stable, the plant must produce no new material that looks like an allergen, and the plant must have (at least) the original nutritional value expected of that food. In essence, it must be the same familiar plant you start with except for the (predicted) new trait encoded by the transgene. And of course the protein encoded by the transgene must be safe—for consumption by humans or animals if it is food or feed, and by nontarget organisms in the environment likely to encounter it. Plants made by traditional plant breeding using "wide crossing" to bring in a desired gene from a distant (weedy or progenitor) relative are more likely to have unexpected properties than are transgenic plants. That is because unwanted and unknown genes will always be linked to the desirable trait sought in the wide cross.

The final problem—one still unsolved in many parts of the world—is that the transgenic plant, once certified safe and functional, must be accepted by consumers. Here, I speak as an aging but fond midwife looking at this adolescent technology that I helped to birth. I find that we are now facing a new kind of challenge, one on which all of the science discussed here seems to have surprisingly little impact.

Many consumers oppose transgenic plants as something either dangerous or unethical, possibly both. These opponents are not likely to inform themselves about plant biotechnology by reading materials such as you will find assembled between the covers of this book. But many are at least curious about this unknown thing that they oppose. I hope that many of you who read this book will become informed advocates of plant biotechnology. Talk to the curious. Replace suspicion, where you can, with information. Replace doubt with evidence. I do not think, however, that in order to spread trust, it is necessary to teach everyone about this technology. People are busy. They will not expend the time and energy to inform themselves in depth. I think that you only need to convince people that *you* have studied this subject in detail, that you have read this book, that you harbor no bias, and that you think that it is safe and natural, as I believe you will.

I have invested most of my career in developing and exploiting the technology for putting new genes into plants. My greatest hope is to see wide—at least wider—acceptance of transgenic plants by consumers during my lifetime. Transgene integration by plants is a natural phenomenon, so much so that we are still trying to figure out exactly how Mother Nature does it. *Agrobacterium* was a microbial genetic engineer long before I began studying DNA. Plant biotechnology has already made significant and positive environmental contributions, as you will discover in the very first chapter of this book. It has the potential to be a powerful new tool for plant breeders, one that they will surely need in facing the challenges of rapid climate change, flood and drought, global warming, as well as the new pests and diseases that these changes may bring. The years ahead promise to be very challenging and interesting. I think that this book will serve you readers well as you prepare for your various roles in meeting those challenges. Enjoy your travels through these chapters and beyond, and I sincerely hope that your journey may turn out to be as interesting and rewarding as mine has been.

Mary-Dell Chilton
Syngenta Biotechnology
Research Triangle Park, North Carolina

CONTRIBUTORS

Monica Alandete-Saez, Public Intellectual Property Resource for Agriculture, Department of Plant Sciences, University of California, Davis, California

Detlef Bartsch, Federal Office of Consumer Protection and Food Safety, Berlin, Germany

Alan B. Bennett, Public Intellectual Property Resource for Agriculture, Department of Plant Sciences, University of California, Davis, California

Sara Boettiger, Public Intellectual Property Resource for Agriculture, Department of Plant Sciences, University of California, Davis, California

Graham Brookes, PG Economics Ltd, Frampton, Dorchester, UK

Vinitha Cardoza, BASF Plant Science LP, Research Triangle Park, North Carolina

Cecilia Chi-Ham, Public Intellectual Property Resource for Agriculture, Department of Plant Sciences, University of California, Davis, California; HM Clause, Inc., Davis, California

Elroy R. Cober, Agriculture and Agri-Food Canada, Ottawa, Canada

Mark D. Curtis, Institute of Plant Biology, University of Zurich, Zurich, Switzerland

John J. Finer, Department of Horticulture and Crop Science, OARDC/The Ohio State University, Wooster, Ohio

Alison K. Flynn, Veterinary Medical Center, University of Florida, Gainesville, Florida

Maria Gallo, Molecular Biosciences and Bioengineering Department, University of Hawaii at Mānoa, Honolulu, Hawaii

Achim Gathmann, Federal Office of Consumer Protection and Food Safety, Berlin, Germany

Glenda E. Gillaspy, Department of Biochemistry, Virginia Tech, Blacksburg, Virginia

Gregory Graff, Department of Agricultural & Resource Economics, Colorado State University, Fort Collins, Colorado

Matthew D. Halfhill, Department of Biology, Saint Ambrose University, Davenport, Iowa

Kenneth L. Korth, Department of Plant Pathology, University of Arkansas, Fayetteville, Arkansas

Wusheng Liu, Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

David G.J. Mann, Dow AgroSciences, Indianapolis, Indiana

Alan McHughen, Department of Botany and Plant Sciences, University of California, Riverside, California

Brian Miki, Agriculture and Agri-Food Canada, Ottawa, Canada

Douglas Powell, Brisbane, Australia

Christiane Saeglitz, Biotechnology, Bioeconomy, Health Research, Project Management, Forschungszentrum Jülich GmbH, Jülich, Germany

Arti Sinha, Wayu Health, Gurgaon, India

C. Neal Stewart, Jr., Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

Nicholas A. Tinker, Agriculture and Agri-Food Canada, Ottawa, Canada

Jennifer Trumbo, Department of Nutrition, University of Tennessee, Knoxville, Tennessee

Suzanne I. Warwick, Agriculture and Agri-Food Canada, Eastern Cereal and Oilseeds Research Centre, Ottawa, Canada

PREFACE

I vividly recall having a series of conversations back in the mid-1990s with "older" plant biotechnologists. These were the seasoned veterans who'd been on the cutting edge of figuring out how to make transgenic plants and how they might partially solve some critical problems in agriculture. They had been through the long days, weeks, months, and years of making genetically engineered commercial crops a reality as the middle of that decade saw the first commercial products hit the market. These scientists had worked out the basic science on how to produce recombinant DNA; genetically engineer the novel DNA sequences into plant cells; and then recover, for the first time, genetically engineered crops. They had witnessed challenge after challenge in the lab. They'd plodded through failures—many failures—and then, finally, success! After the promising transgenic crop lines had been produced, then came the arduous process of plant breeding, which was needed to move the useful traits into agronomic varieties that farmers would want to grow. Then came the field testing, seed production, and then…let's not forget about all the regulatory approvals. Each step was like those taken by a toddler. It was all new ground. The difference between walking and falling down was measured in millimeters. And the baby put one foot in front of the other, often with great pauses to regain balance. Finally, the faithful day would arrive when the genetically engineered seed would be planted and bear fruit in farmers' fields. And there we were.

It wasn't a shock in the mid-1990s when these scientists expressed to me their feelings that went something like, "all the really fun stuff has already been done." I was still a pretty young scientist at the time, and so who was I to question their insights? These insights from giants who stood on the shoulders of giants? So, in these awestruck moments, I asked polite questions, listened to their stories, and like a fawning fan I would muster an occasional "cool!" To be honest, their words and attitudes took a little wind out of my sails after I went back to my own little lab and office. From their perspective, indeed, the big challenges of moving those first molecules from idea to seed could never be matched again. But still, I thought about the future of the field and plodded along with my own ideas and research. I wanted to make the world a better place and believed that we could innovate with plant biotechnology—even, maybe, despite the assertion that all the coolest and most fun stuff had already been done. So I thought.

When we fast-forward about 10 years later, I thought it would be a fun project to put together a plant biotech textbook to support the course I'd offered to teach. The product of all the fun would be what became the first edition of the title in your hands. As that book came together, I sometimes thought about what I'd been told by these sages. The content of the text in the book, it seemed, mostly consisted of the tried and true technologies that were used in making those first engineered plants. There were also stories told of the glory days by scientists who penned their "Life boxes" in the book. After a while, however, I noticed that the first edition was starting to be somewhat dated itself. There were now new DNA sequencing technologies. There were new analytical techniques. New genome editing tools and synthetic biology tools had been invented and it was clear they would have an impact on plants. Computers had also changed what could be done and the speed tasks could be performed. So I embarked on updating the book and the second edition took shape.

Sometime in the last year or so, while working on the book, it really started to hit me, and has since pounded me like a John Henry sledgehammer on railroad spikes: those good old days were not the best days of plant biotechnology after all. The best and most fun stuff has not been done yet.

Yes, of course, a baby only learns to walk once, but now plant biotechnologists could sprint. It became clear that genome editing tools could allow biotechnologists to reconfigure existing genes in plants in ways never imagined by the early pioneers of biotechnology. Recently, a chromosome has been totally synthesized and installed into yeast—how long would it be before whole new entire pathways could be installed into plants to enable them to do things not even thought possible in the good old days? I have become convinced that the most intriguing and exciting days in plant biology and biotechnology are to be ushered in as computationally enabled genetics matures and becomes widely utilized. Crop productivity will continue to be improved using new innovations. Increased yield will feed more people with more nutritious food. And the readers of this book will be the ones to usher in the next wave of innovation. That is best and most fun part for me right now—making the future reality.

The second edition contains all updated chapters and new chapters in systems and synthetic biology. The "Life box" profiles of the plant biologists and biotechnologists who have made a difference in the field have been updated and the number of scientists who are profiled has been expanded. The lecture slides for open access to instructors and students remain at http://plantsciences.utk.edu/pbg/, and these are updated each time I teach the class. Feel free to offer any suggestions or slides of your own that I could use to update this resource.

I'm very grateful to the chapter authors and Life Box authors—both carried over from the first edition of the book—and the new ones. Thanks to my lab crew for their patience during the preparation of the book. I'm particularly indebted to Jennifer Hinds at the University of Tennessee. Jennifer did so much work on the book, I can't begin make a list of her contributions. This much is certain: without Jennifer, there would be no second edition of the book. Thanks, Jennifer! You're awesome!!

C. Neal Stewart
Knoxville, Tennessee
June 21, 2015

CHAPTER 1

The Impact of Biotechnology on Plant Agriculture

GRAHAM BROOKES

PG Economics Ltd, Frampton, Dorchester, UK

1.0. CHAPTER SUMMARY AND OBJECTIVES

1.0.1. Summary

Since the first stably transgenic plant produced in the early 1980s and the first commercialized transgenic plant in 1994, biotechnology has revolutionized plant agriculture. In the United States, between 80 and 90% of the maize (corn), soybean, cotton, and canola crops are transgenic for insect resistance, herbicide resistance, or both. Biotechnology has been the most rapidly adopted technology in the history of agriculture and continues to expand in much of the developed and developing world.

1.0.2. Discussion Questions

1. What biotechnology crops are grown and where?
2. Why do farmers use biotech crops?
3. How has the adoption of plant biotechnology impacted the environment?

1.1. INTRODUCTION

The technology of genetic modification (GM, also stands for "genetically modified"), which consists of genetic engineering and also known as genetic transformation, has now been utilized globally on a widespread commercial basis for 18 years; and by 2012, 17.3 million farmers in 28 countries had planted 160 million hectares of crops using this technology. These milestones provide an opportunity to critically assess the impact of this technology on global agriculture. This chapter therefore examines specific global socioeconomic impacts on farm income and environmental impacts with respect to pesticide usage and greenhouse gas (GHG) emissions of the technology. Further details can be found in Brookes and Barfoot (2014a, b).

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

1.2. CULTIVATION OF BIOTECHNOLOGY (GM) CROPS

Although the first commercial GM crops were planted in 1994 (tomatoes), 1996 was the first year in which a significant area of crops containing GM traits were planted (1.66 million hectares). Since then, there has been a dramatic increase in plantings, and by 2012 the global planted area reached over 160.4 million hectares.

Almost all of the global GM crop area derives from soybean, maize (corn), cotton, and canola (Fig. 1.1). In 2012, GM soybean accounted for the largest share (49%) of total GM crop cultivation, followed by maize (32%), cotton (14%), and canola (5%). In terms of the share of total global plantings to these four crops accounted for by GM crops, GM traits accounted for a majority of soybean grown (73%) in 2012 (i.e., non-GM soybean accounted for 27% of global soybean acreage in 2012). For the other three main crops, the GM shares in 2012 of total crop production were 29% for maize, 59% for cotton, and 26% for canola (i.e., the majority of global plantings of maize and canola continued to be non-GM in 2012). The trend in plantings of GM crops (by crop) from 1996 to 2012 is shown in Figure 1.2. In terms of the type of biotechnology trait planted, Figure 1.3 shows that GM herbicide-tolerant soybeans dominate, accounting for 38% of the total, followed by insect-resistant (largely Bt) maize, herbicide-tolerant maize, and insect-resistant cotton with respective shares of 26, 19, and 11%. It is worth noting that the total number of plantings by trait produces a higher global planted area (209.2 million hectares) than the global area by crop (160.4 million hectares) because of the planting of some crops containing the stacked traits of herbicide tolerance and insect resistance (e.g., a single plant with two biotech traits).

In total, GM herbicide-tolerant (GM HT) crops account for 63%, and GM insect-resistant (GM IR) crops account for 37% of global plantings. Finally, looking at where biotech crops have been grown, the United States had the largest share of global GM crop plantings in 2012

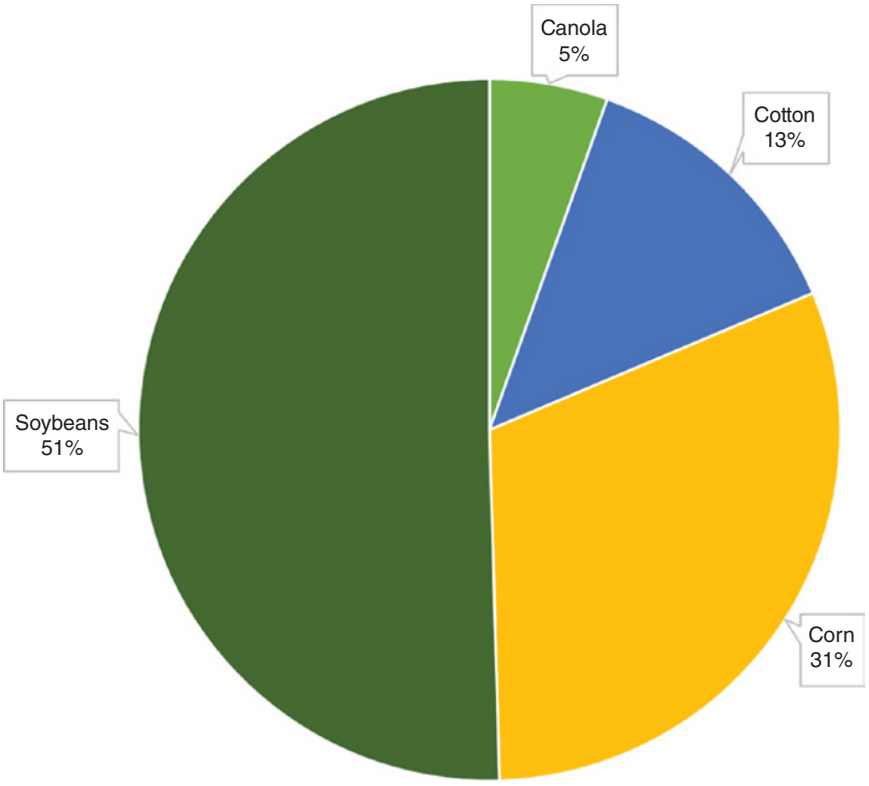


Figure 1.1. Global GM crop plantings in 2012 by crop (base area: 160.4 million hectare). (*Sources*: ISAAA, Canola Council of Canada, CropLife Canada, USDA, CSIRO, ArgenBio.)

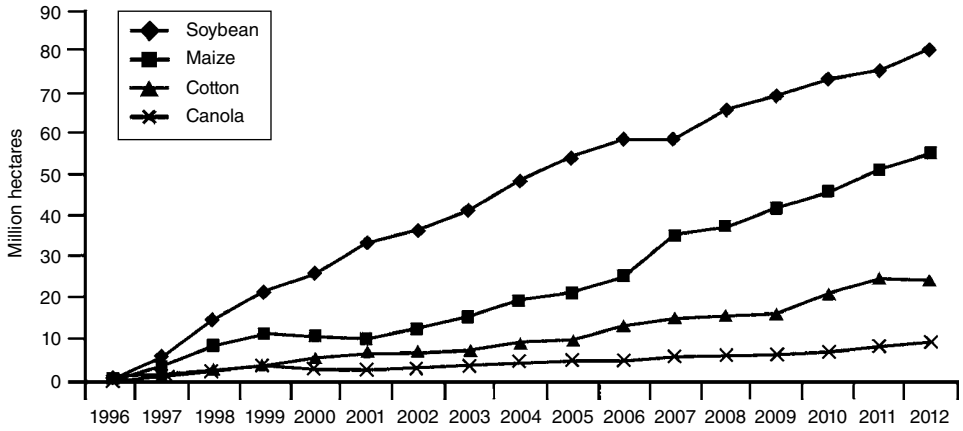


Figure 1.2. Global GM crop plantings by crop 1996–2012. (*Sources*: ISAAA, Canola Council of Canada, CropLife Canada, USDA, CSIRO, ArgenBio.)

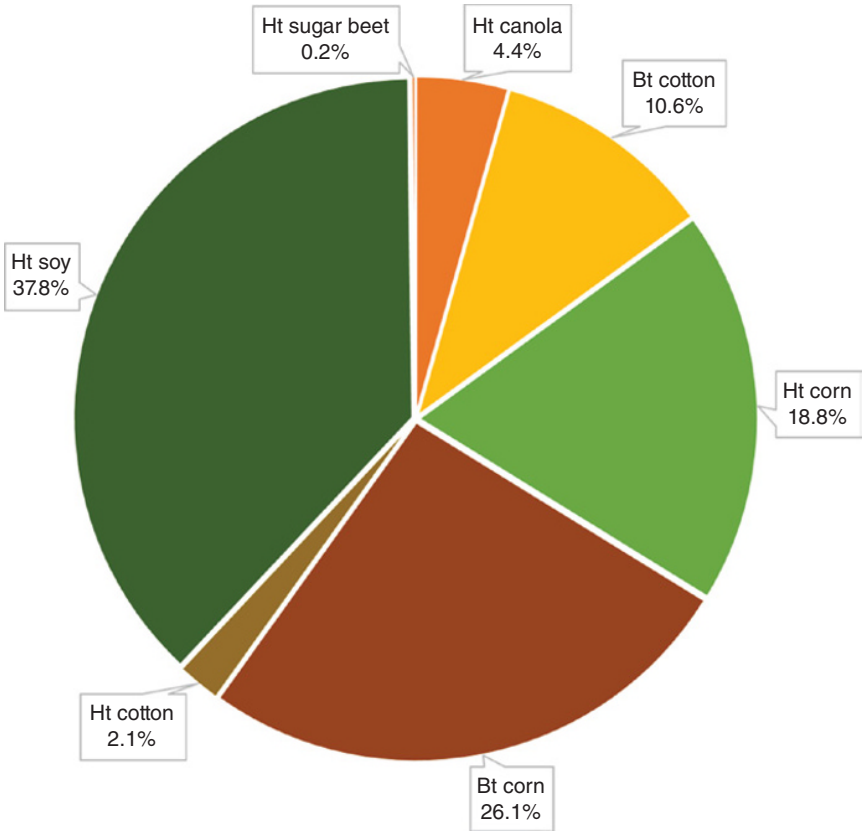


Figure 1.3. Global GM crop plantings by main trait and crop: 2012. (*Sources*: Various, including ISAAA, Canola Council of Canada, CropLife Canada, USDA, CSIRO, ArgenBio.)

(40%: 64.1 million hectares), followed by Brazil (37.2 million hectares: 23% of the global total) and Argentina (14%: 23.1 million hectares). The other main countries planting GM crops in 2012 were India, Canada, and China (Fig. 1.4). In 2012, there were also additional GM crop plantings of papaya (395 hectares), squash (2000 hectares), alfalfa (425,000 hectares), and sugar

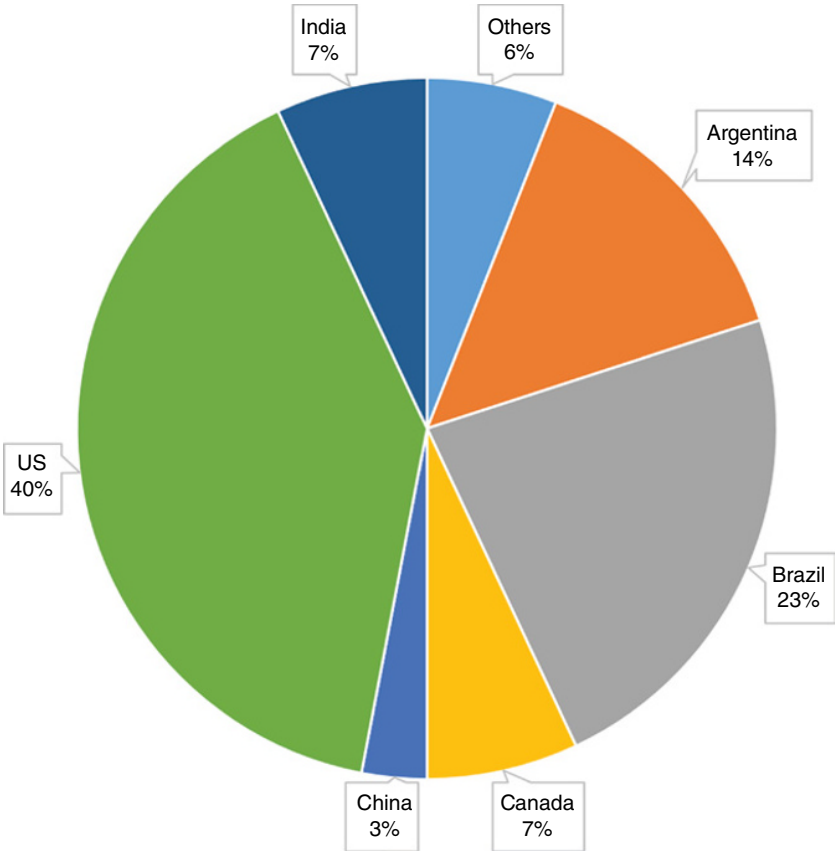


Figure 1.4. Global GM crop plantings 2012 by country. (*Sources*: ISAAA, Canola Council of Canada, CropLife Canada, USDA, CSIRO, ArgenBio.)

beet (490,000 hectares) in the United States, of papaya (5000 hectares) in China and of sugar beet (13,500 hectares) in Canada.

1.3. WHY FARMERS USE BIOTECH CROPS

The primary driver of adoption among farmers (both large commercial and small-scale subsistence) has been the positive impact on farm income. The adoption of biotechnology has had a very positive impact on farm income derived mainly from a combination of enhanced productivity and efficiency gains (Table 1.1). In 2012, the direct global farm income benefit from GM crops was $18.8 billion. This is equivalent to having added 5.6% to the value of global production of the four main crops of soybean, maize, canola, and cotton, a substantial impact. Since 1996, worldwide farm incomes have increased by $116.6 billion, directly because of the adoption of GM crop technology.

The largest gains in farm income in 2012 have arisen in the maize sector, largely from yield gains. The $6.7 billion additional income generated by GM IR maize in 2012 has been equivalent to adding 6.6% to the value of the crop in the GM crop-growing countries, or adding the equivalent of 3% to the $226 billion value of the global maize crop in 2012. Cumulatively since 1996, GM IR technology has added $32.3 billion to the income of global maize farmers.

Substantial gains have also arisen in the cotton sector through a combination of higher yields and lower costs. In 2012, cotton farm income levels in the GM-adopting countries increased by

TABLE 1.1. Global Farm Income Benefits from Growing GM Crops 1996–2012 (Million US $)

Trait	Increase in farm income 2012	Increase in farm income 1996–2012	Farm income benefit in 2012 as percentage of total value of production of these crops in GM adopting countries	Farm income benefit in 2012 as percentage of total value of global production of crop
GM herbicide-tolerant soybeans	4,797.9	37,008.6	4.4	4.0
GM herbicide-tolerant maize	1,197.9	5,414.7	1.2	0.5
GM herbicide-tolerant cotton	147.2	1,371.6	0.4	0.3
GM herbicide-tolerant canola	481.0	3,664.4	4.9	1.3
GM insect-resistant maize	6,727.8	32,317.2	6.6	3.0
GM insect-resistant cotton	5,331.3	36,317.2	13.1	11.2
Others	86.3	496.7	N/A	N/A
Total	18,769.4	116,590.4	6.8	5.6

Notes: All values are nominal. Others = Virus resistant papaya and squash and herbicide-tolerant sugar beet. Totals for the value shares exclude "other crops" (i.e., relate to the four main crops of soybeans, maize, canola, and cotton). Farm income calculations are net farm income changes after inclusion of impacts on yield, crop quality, and key variable costs of production (e.g., payment of seed premia, impact on crop protection expenditure). N/A = not applicable.

$5.5 billion; and since 1996, the sector has benefited from an additional $37.7 billion. The 2012 income gains are equivalent to adding 13.5% to the value of the cotton crop in these countries, or 11.5% to the $47 billion value of total global cotton production. This is a substantial increase in value-added terms for two new cotton seed technologies.

Significant increases to farm incomes have also resulted in the soybean and canola sectors. The GM HT technology in soybeans has boosted farm incomes by $4.8 billion in 2012, and since 1996 has delivered over $37 billion of extra farm income. In the canola sector (largely North American) an additional $3.66 billion has been generated (1996–2012).

Overall, the economic gains derived from planting GM crops have been of two main types: (a) increased yields (associated mostly with GM IR technology) and (b) reduced costs of production derived from less expenditure on crop protection (insecticides and herbicides) products and fuel.

Table 1.2 summarizes farm income impacts in key GM-adopting countries highlighting the important farm income benefit arising from GM HT soybeans in South America (Argentina, Bolivia, Brazil, Paraguay, and Uruguay), GM IR cotton in China and India, and a range of GM cultivars in the United States. It also illustrates the growing level of farm income benefits being obtained in South Africa, the Philippines, Mexico, and Colombia from planting GM crops.

In terms of the division of the economic benefits, it is interesting to note that farmers in developing countries derived in 2012 (46.2%) relative to farmers in developed countries (Table 1.3). The vast majority of these income gains for developing country farmers have been from GM IR cotton and GM HT soybean.[1]

[1]The author acknowledges that the classification of different countries into "developing" or "developed" status affects the distribution of benefits between these two categories of country. The definition used here is consistent with the definition used by others, including the International Service for the Acquisition of Agri-Biotech Applications (ISAAA) (see the review by James (2012)].

TABLE 1.2. GM Crop Farm Income Benefits During 1996–2012 in Selected Countries (Million US $)

	GM HT soybeans	GM HT maize	GM HT cotton	GM HT canola	GM IR maize	GM IR cotton	Total
United States	16,668.7	3752.3	975.8	268.3	26,375.9	4,046.7	52,087.7
Argentina	13,738.5	766.7	107.0	N/A	495.2	456.4	15,563.8
Brazil	4,825.6	703.4	92.5	N/A	2,761.7	13.3	8,396.5
Paraguay	828	N/A	N/A	N/A	N/A	N/A	828.0
Canada	358	81.3	N/A	3368.8	1,042.9	N/A	4,851.0
South Africa	9.1	4.1	3.2	N/A	1,100.6	34.2	1,151.2
China	N/A	N/A	N/A	N/A	N/A	15,270.4	15,270.4
India	N/A	N/A	N/A	N/A	N/A	14,557.1	14,557.1
Australia	N/A	N/A	78.6	27.3	N/A	659.6	765.5
Mexico	5.0	N/A	96.4	N/A	N/A	136.6	238.0
Philippines	N/A	104.7	N/A	N/A	273.6	N/A	378.3
Romania	44.6	N/A	N/A	N/A	N/A	N/A	44.6
Uruguay	103.8	N/A	N/A	N/A	17.6	N/A	121.4
Spain	N/A	N/A	N/A	N/A	176.3	N/A	176.3
Other EU	N/A	N/A	N/A	N/A	18.8	N/A	18.8
Colombia	N/A	1.7	18.1	N/A	47.4	15.4	826.6
Bolivia	432.2	N/A	N/A	N/A	N/A	N/A	432.2
Burma	N/A	N/A	N/A	N/A	N/A	215.4	215.4
Pakistan	N/A	N/A	N/A	N/A	N/A	725.1	725.1
Burkina Faso	N/A	N/A	N/A	N/A	N/A	186.9	186.9
Honduras	N/A	N/A	N/A	N/A	6.9	N/A	6.9

Notes: All values are nominal. Farm income calculations are net farm income changes after inclusion of impacts on yield, crop quality, and key variable costs of production (e.g., payment of seed premia, impact on crop protection expenditure). N/A = not applicable. US total figure also includes $491 million for other crops/traits (not included in the table). Also not included in the table is $5.5 million extra farm income from GM HT sugar beet in Canada.

TABLE 1.3. GM Crop Farm Income Benefits, 2012: Developing Versus Developed Countries (Million US $)

	Developed	Developing
GM HT soybeans	2,955.4	1842.5
GM HT maize	654.0	543.9
GM HT cotton	71.4	75.8
GM HT canola	481.0	0
GM IR maize	5,327.5	1400.3
GM IR cotton	530.7	4800.7
GM virus-resistant papaya and squash and GM HT sugar beet	86.3	0
Total	10,106.3	8663.2

Note: Developing countries = All countries in South America, Mexico, Honduras, Burkina Faso, India, China, the Philippines, and South Africa.

Examination of the cost farmers pay for accessing GM technology relative to the total gains derived shows that across the four main GM crops, the total cost was equal to about 23% of the total farm income gains (Table 1.4). For farmers in developing countries, the total cost is equal to about 21% of total farm income gains, while for farmers in developed countries the cost is about 25% of the total farm income gain. Although circumstances vary between countries, the higher share of total technology gains accounted for by farm income gains in developing countries, relative to the farm income share in developed countries, reflects factors such as weaker provision and enforcement of intellectual property rights in developing countries and the higher average

TABLE 1.4. Cost of Accessing GM Technology Relative to Total Farm Income Benefits (US Millions) 2012

	Tech costs: all farmers	Farm income gain: all farmers	Total benefit of technology to farmers and seed supply chain	Cost of technology: developing countries	Farm income gain: developing countries	Total benefit of technology to farmers and seed supply chain: developing countries
GM HT soy	1528.1	4,797.9	6,326.0	998.7	1842.5	2,841.2
GM HT maize	1059.4	1,197.9	2,257.3	364.5	543.9	908.4
GM HT cotton	295.0	147.2	442.2	22.2	75.8	98.0
GM HT canola	161.2	481.0	642.2	N/A	N/A	N/A
GM IR maize	1800.8	6,727.8	8,528.6	512.3	1400.3	1,912.6
GM IR cotton	720.7	5,331.3	6,052.0	422.7	4800.7	5,223.4
Others	76.2	86.3	162.5	N/A	N/A	N/A
Total	5641.4	18,769.4	24,410.8	2320.4	8663.2	10,983.6

N/A = not applicable. Cost of accessing technology based on the seed premiums paid by farmers for using GM technology relative to its conventional equivalents.

level of farm income gain on a per-hectare basis derived by developing country farmers relative to developed country farmers.

In addition to the tangible and quantifiable impacts on farm profitability presented earlier, there are other important, more intangible (difficult to quantify) impacts of an economic nature. Many studies on the impact of GM crops have identified the factors listed later in the text as being important influences for the adoption of the technology.

1.4. GM'S EFFECTS ON CROP PRODUCTION AND FARMING

Based on the yield impacts used in the direct farm income benefit calculations discussed earlier and taking account of the second soybean crop facilitation in South America, GM crops have added important volumes to global production of maize, cotton, canola, and soybeans since 1996 (Table 1.5).

The GM IR traits, used in maize and cotton, have accounted for 97.1% of the additional maize production and 99.3% of the additional cotton production. Positive yield impacts from the use of this technology have occurred in all user countries (except for GM IR cotton in Australia[2]) when compared to average yields derived from crops using conventional technology (i.e., application of insecticides and seed treatments). The average yield impact across the total area planted to these traits over the 17 years since 1996 has been +10.4% for maize and +16.1% for cotton.

As indicated earlier, the primary impact of GM HT technology has been to provide more cost-effective (less-expensive) and easier weed control, as opposed to improving yields. The improved weed control has, nevertheless, delivered higher yields in some countries. The main source of additional production from this technology has been via the facilitation of no-tillage production system, shortening the production cycle and how it has enabled many farmers in South America to plant a crop of soybeans immediately after a wheat crop in the same growing season. This second crop, additional to traditional soybean production, has added 114.3 million tonnes to soybean production in Argentina and Paraguay between 1996 and 2012 (accounting for 93.5% of the total GM-related additional soybean production).

[2]This reflects the levels of *Heliothis*/*Helicoverpa* (boll and bud worm pests) pest control previously obtained with intensive insecticide use. The main benefit and reason for adoption of this technology in Australia has arisen from significant cost savings (on insecticides) and the associated environmental gains from reduced insecticide use.

TABLE 1.5. Additional Crop Production Arising from Positive Yield Effects of GM Crops

	1996–2012 additional production (million tonnes)	2012 additional production (million tonnes)
Soybeans	122.3	12.0
Maize	231.4	34.1
Cotton	18.2	2.4
Canola	6.6	0.4
Sugar beet	0.6	0.15

Note: GM HT sugar beet has been commercialized only in the United States and Canada since 2008.

1.5. HOW THE ADOPTION OF PLANT BIOTECHNOLOGY HAS IMPACTED THE ENVIRONMENT

Two key aspects of environmental impact of biotech crops examined later are decreased insecticide and herbicide use, and the impact on carbon emissions and soil conservation.

1.5.1. Environmental Impacts from Changes in Insecticide and Herbicide Use

Usually, changes in pesticide use with GM crops have traditionally been presented in terms of the volume (quantity) of pesticide applied. While comparisons of total pesticide volume used in GM and non-GM crop production systems can be a useful indicator of environmental impacts, it is an imperfect measure because it does not account for differences in the specific pest control programs used in GM and non-GM cropping systems. For example, different specific chemical products used in GM versus conventional crop systems, differences in the rate of pesticides used for efficacy, and differences in the environmental characteristics (mobility, persistence, etc.) are masked in general comparisons of total pesticide volumes used.

To provide a more robust measurement of the environmental impact of GM crops, the analysis presented in the following text includes an assessment of both pesticide active-ingredient use and the specific pesticides used via an indicator known as the environmental impact quotient (EIQ). This universal indicator, developed by Kovach et al. (1992) and updated annually, effectively integrates the various environmental impacts of individual pesticides into a single field value per hectare. This index provides a more balanced assessment of the impact of GM crops on the environment as it draws on all of the key toxicity and environmental exposure data related to individual products, as applicable to impacts on farmworkers, consumers, and ecology, and provides a consistent and comprehensive measure of environmental impact. Readers should, however, note that the EIQ is an indicator only and, therefore, does not account for all environmental issues and impacts.

The EIQ value is multiplied by the amount of pesticide active ingredient (AI) used per hectare to produce a field EIQ value. For example, the EIQ rating for glyphosate is 15.3. By using this rating multiplied by the amount of glyphosate used per hectare (e.g., a hypothetical example of 1.1 kg applied per hectare), the field EIQ value for glyphosate would be equivalent to 16.83/hectare. In comparison, the field EIQ/hectare value for a commonly used herbicide on corn crops (atrazine) is 22.9/hectare.

The EIQ indicator is therefore used for comparison of the field EIQ/hectare values for conventional versus GM crop production systems, with the total environmental impact or load of each system, a direct function of respective field EIQ/hectare values, and the area planted to each type of production (GM vs. non-GM).

The EIQ methodology is used in the following to calculate and compare typical EIQ values for conventional and GM crops and then aggregate these values to a national level. The level of pesticide

use in the respective areas planted for conventional and GM crops in each year was compared with the level of pesticide use that probably would otherwise have occurred if the whole crop, in each year, had been produced using conventional technology (based on the knowledge of crop advisers). This approach addresses gaps in the availability of herbicide or insecticide usage data in most countries and differentiates between GM and conventional crops. Additionally, it allows for comparisons between GM and non-GM cropping systems when GM accounts for a large proportion of the total crop planted area. For example, in the case of soybean in several countries, GM represents over 60% of the total soybean crop planted area. It is not reasonable to compare the production practices of these two groups as the remaining non-GM adopters might be farmers in a region characterized by below-average weed or pest pressures or with a tradition of less intensive production systems, and hence, below-average pesticide use.

GM crops have contributed to a significant reduction in the global environmental impact of production agriculture (Table 1.6). Since 1996, the use of pesticides was reduced by 503 million kg of AI, constituting an 8.8% reduction, and the overall environmental impact associated with pesticide use on these crops was reduced by 18.7%. In absolute terms, the largest environmental gain has been associated with the adoption of GM IR technology. GM IR cotton has contributed a 25.6% reduction in the volume of AI used and a 28.2% reduction in the EIQ indicator (1996–2012) due to the significant reduction in insecticide use that the technology has facilitated, in what has traditionally been an intensive user of insecticides. Similarly, the use of GM IR technology in maize has led to important reductions in insecticide use, with associated environmental benefits.

The volume of herbicides used in GM maize crops also decreased by 203 million kg (1996–2012), a 9.8% reduction, whilst the overall environmental impact associated with herbicide use on these crops decreased by a significantly larger 13.3%. This highlights the switch in herbicides used with most GM HT crops to AIs with a more environmentally benign profile than the ones generally used on conventional crops.

TABLE 1.6. Impact of Changes in the Use of Herbicides and Insecticides from Global Cultivation of GM Crops, Including Environmental Impact Quotient (EIQ), 1996–2012

Trait	Change in mass of active ingredient used (million kg)	Change in field EIQ (in terms of million field EIQ/ hectare units)	Percentage change in AI use on GM crops	Percentage change in environmental impact associated with herbicide and insecticide use on GM crops	Area GM trait 2012 (million hectare)
GM herbicide-tolerant soybeans	−4.7	−6,654	−0.2	−15.0	79.1
GM herbicide-tolerant maize	−203.2	−6,025	−9.8	−13.3	38.5
GM herbicide-tolerant canola	−15.0	−509	−16.7	−26.6	8.6
GM herbicide-tolerant cotton	−18.3	−460	−6.6	−9.0	4.4
GM insect-resistant maize	−57.6	−2,215	−47.9	−45.1	42.3
GM insect-resistant cotton	−205.4	−9,256	−25.6	−28.2	22.1
GM herbicide-tolerant sugar beet	+1.3	−2	+29.3	−2.0	0.51
Totals	−503.1	−25,121	−8.8	−18.7	

TABLE 1.7. Changes in Environmental Impact Quotient (EIQ) form GM Crops and Associated Changes in Associated Insecticide and Herbicide Use in 2012: Developing versus Developed Countries

	Change in field EIQ (in terms of million field EIQ/hectare units): developed countries	Change in field EIQ (in terms of million field EIQ/hectare units): developing countries
GM HT soybeans	−4,773.9	−1,880.2
GM HT maize	−5,585.9	−438.8
GM HT cotton	−351.0	−109.3
GM HT canola	−509.1	0
GM IR maize	−1,574.4	−640.8
GM IR cotton	−805.5	−8,451.0
GM HT sugar beet	−2	0
Total	−13,601.8	−11,520.1

Important environmental gains have also arisen in the soybean and canola sectors. In the soybean sector, herbicide use decreased by 4.7 million kg (1996–2012) and the associated environmental impact of herbicide use on this crop area decreased, from a switch to more environmentally benign herbicides (−15%). In the canola sector, farmers reduced herbicide use by 15 million kg (a 16.7% reduction) and the associated environmental impact of herbicide use on this crop area fell by 26.6% (from switching to more environmentally benign herbicides).

In terms of the division of the environmental benefits associated with less insecticide and herbicide use for farmers in developed countries relative to farmers in developing countries, Table 1.7 shows a 54 : 46% split of the environmental benefits (1996–2012), respectively, in developed (54%) and developing countries (46%). About three-quarters (73%) of the environmental gains in developing countries have been from the use of GM IR cotton.

It should, however, be noted that in some regions where GM HT crops have been widely grown, some farmers have relied too much on the use of single herbicides, such as glyphosate, to manage weeds in GM HT crops and this has contributed to the evolution and spread of weed resistance. There are currently 31 weed species recognized as exhibiting resistance to glyphosate worldwide, of which several are not associated with glyphosate-tolerant crops (www.weedscience.org). For example, there are currently 14 weeds recognized in the United States as exhibiting resistance to glyphosate, of which two are not associated with glyphosate tolerant crops. In the United States, the affected area is currently within a range of 15–40% of the total area annually devoted to maize, cotton, canola, soybeans, and sugar beet (the crops in which GM HT technology is used).

In recent years, there has also been a growing consensus among weed scientists of a need for changes in the weed management programs in GM HT crops, because of the apparent increase of evolution glyphosate-resistant weeds. Growers of GM HT crops are increasingly being advised to be more proactive and include other herbicides (with different and complementary modes of action) in combination with glyphosate in their integrated weed management systems, even where instances of weed resistance to glyphosate have not been found.

This proactive, diversified approach to weed management is the principal strategy for avoiding the emergence of HR weeds in GM HT crops. It is also the main way of tackling weed resistance in conventional crops. A proactive weed management program also generally requires using less herbicide, has a better environmental profile, and is more economical than a reactive weed management program.

At the macrolevel, the adoption of both reactive and proactive weed management programs in GM HT crops has already begun to influence the mix, total amount and overall environmental profile of herbicides applied to GM HT soybeans, cotton, maize, and canola, and this is reflected in the data presented in this chapter.

1.5.2. Impact on GHG Emissions

The reduction in the levels of GHG emissions from GM crops are from the following two principal sources:

1. GM crops contribute to a reduction in fuel use from less frequent herbicide or insecticide applications and a reduction in the energy use in soil cultivation. For example, Lazarus (2012) estimated that one pesticide spray application uses 0.841 of fuel per hectare, which is equivalent to 2.24 kg/hectare of carbon dioxide emissions. In this analysis, we used the conservative assumption that only GM IR crops reduced spray applications and ultimately GHG emissions. In addition to the reduction in the number of herbicide applications, there has been a shift from conventional tillage to no-/reduced tillage (NT) and herbicide-based weed control systems, which has had a marked effect on tractor fuel consumption. The GM HT crop where this is most evident is GM HT soybean and where the GM HT soybean and maize rotation is widely practiced, for example in the United States. Here, adoption of the technology has made an important contribution to facilitating the adoption of NT farming (CTIC 2002, American Soybean Association 2001). Before the introduction of GM HT soybean cultivars, NT systems were practiced by some farmers using a number of herbicides and with varying degrees of success. The opportunity for growers to control weeds with a nonresidual foliar herbicide as a "burndown" preseeding treatment, followed by a postemergent treatment when the soybean crop became established, has made the NT system more reliable, technically viable, and commercially attractive. These technical advantages, combined with the cost advantages, have contributed to the rapid adoption of GM HT cultivars and the near-doubling of the NT soybean area in the United States (and also a ≥sevenfold increase in Argentina). In both countries, GM HT soybean crops are estimated to account for 95% of the NT soybean crop area. Substantial growth in NT production systems has also occurred in Canada, where the NT canola area increased from 0.8 to 8 million hectares a (equal to about 90% of the total canola area) between 1996 and 2012 (95% of the NT canola area is planted with GM HT cultivars). The area planted to NT in the US cotton crop increased from 0.2 to 1 million hectare 1996–2005 (86% of which is planted to GM HT cultivars), although the NT cotton area has not risen above about 25% of the total crop. The fuel savings used in this chapter are drawn from a review of literature including Jasa (2002), CTIC (2002), University of Illinois (2006), USDA Energy Estimator (USDA 2013b), Reeder (2010), and the USDA Comet-VR model (USDA 2013a). It is assumed that the adoption of NT farming systems in soybean production reduces cultivation and seedbed preparation fuel usage by 27.12 l/hectare compared with traditional conventional tillage and in the case of RT (mulch till) cultivation by 10.39 l/hectare. In the case of maize, NT results in a saving of 24.41 l/hectare and 7.52 l/hectare in the case of RT compared with conventional intensive tillage. These are conservative estimates and are in line with the USDA Energy Estimator for soybeans and maize.

 The adoption of NT and RT systems in respect of fuel use therefore results in reductions of carbon dioxide emissions of 72.41 kg/hectare and 27.74 kg/hectare respectively for soybeans and 65.17 kg/hectare and 20.08 kg/hectare for maize.
2. The use of NT[3] farming systems that utilize less plowing increases the amount of organic carbon in the form of crop residue that is stored or sequestered in the soil. This carbon

[3] NT farming means that the ground is not plowed at all, while reduced tillage means that the ground is disturbed less than it would be with traditional tillage systems. For example, under an NT farming system, soybean seeds are planted through the organic material that is left over from a previous crop such as corn, cotton, or wheat. NT systems also significantly reduce soil erosion, and hence deliver both additional economic benefits to farmers, enabling them to cultivate land that might otherwise be of limited value and environmental benefits from the avoidance of loss of flora, fauna, and landscape features.

sequestration reduces carbon dioxide emissions to the environment. Rates of carbon sequestration have been calculated for cropping systems using normal tillage and reduced tillage, and these were incorporated in our analysis on how GM crop adoption has significantly facilitated the increase in carbon sequestration, ultimately reducing the release of CO_2 into the atmosphere. Of course, the amount of carbon sequestered varies by soil type, cropping system, and ecoregion.

Drawing on the literature and models referred to earlier, the analysis presented in the following text has several assumptions by country and crop. For the United States, the soil carbon sequestered by tillage system for maize in continuous rotation with soybeans is assumed to be a net sink of 250 kg of carbon/hectare/year based on NT systems store 251 kg of carbon/hectare/year, RT systems store 75 kg of carbon/hectare/, and CT systems store 1 kg of carbon/hectare/year. For the United States, the soil carbon sequestered by tillage system for soybeans in a continuous rotation with maize is assumed to be a net sink of 100 kg of carbon/hectare/year based on NT systems release 45 kg of carbon/hectare/year, RT systems release 115 kg of carbon/hectare/year, and CT systems release 145 kg of carbon/hectare/year.

For Argentina and Brazil, soil carbon retention is 275 kg carbon/hectare/year for NT soybean cropping and CT systems release 25 kg carbon/hectare/year (a difference of 300 kg carbon/hectare/year).

Table 1.8 summarizes the impact on GHG emissions associated with the planting of GM crops between 1996 and 2012. In 2012, the permanent CO_2 savings from reduced fuel use associated with GM crops was 2111 million kg. This is equivalent to removing 900,000 cars from the road for a year.

TABLE 1.8. Impact of GM Crops on Carbon Sequestration Impact in 2012; Car Equivalents

Crop/trait/country	Permanent carbon dioxide savings arising from reduced fuel use (million kg of carbon dioxide)	Permanent fuel savings: as average family car equivalents removed from the road for a year ('000s)	Potential additional soil carbon sequestration savings (million kg of carbon dioxide)	Soil carbon sequestration savings: as average family car equivalents removed from the road for a year ('000s)
US: GM HT soybeans	210	93	1,070	475
Argentina: GM HT soybeans	736	327	11,186	4,972
Brazil GM HT soybeans	394	175	5,985	2,660
Bolivia, Paraguay, Uruguay: GM HT soybeans	156	69	2,365	1,051
Canada: GM HT canola	203	90	1,024	455
US: GM HT corn	210	93	2,983	1,326
Global GM IR cotton	45	20	0	0
Brazil IR corn	157	69	0	0
Total	2,111	936	24,613	10,939

Notes: Assumption: an average family car produces 150 g of carbon dioxide per km. A car does an average of 15,000 km/year and therefore produces 2,250 kg of carbon dioxide/year.

The additional soil carbon sequestration gains resulting from reduced tillage with GM crops accounted for a reduction of 24,613 million kg of CO_2 emissions in 2012. This is equivalent to removing nearly 10.9 million cars from the roads per year. In total, the carbon savings from reduced fuel use and soil carbon sequestration in 2012 were equal to removing 11.88 million cars from the road (equal to 41% of all registered cars in the United Kingdom).

1.6. CONCLUSIONS

Crop biotechnology has, to date, delivered several specific agronomic traits that have overcome a number of production constraints for many farmers. This has resulted in improved productivity and profitability for the 17.3 million adopting farmers who have applied the technology to 160 million hectares in 2012.

During the past 17 years, this technology has made important positive socioeconomic and environmental contributions. These have arisen even though only a limited range of GM agronomic traits have so far been commercialized, in a small range of crops.

Crop biotechnology has delivered economic and environmental gains through a combination of their inherent technical advances and the role of the technology in the facilitation and evolution of more cost effective and environment-friendly farming practices. More specifically the following:

The gains from the GM IR traits have mostly been delivered directly from the technology (yield improvements, reduced production risk and decreased use of insecticides). Thus, farmers (mostly in developing countries) have been able to both improve their productivity and economic returns, whilst also practicing more environment-friendly farming methods;

The gains from GM HT traits have come from a combination of direct benefits (mostly cost reductions to the farmer) and the facilitation of changes in farming systems. Thus, GM HT technology (especially in soybeans) has played an important role in enabling farmers to capitalize on the availability of a low cost, broad-spectrum herbicide (glyphosate) and, in turn, facilitated the move away from conventional to low-/no-tillage production systems in both North and South America. This change in production system has made additional positive economic contributions to farmers (and the wider economy) and delivered important environmental benefits, notably reduced levels of GHG emissions (from reduced tractor fuel use and additional soil carbon sequestration).

Both IR and HT traits have made important contributions to increasing world production levels of soybeans, corn, cotton, and canola.

In relation to GM HT crops, however, overreliance on the use of glyphosate by some farmers, in some regions, has contributed to the evolution and spread of HR weeds. As a result, farmers are increasingly adopting a mix of reactive and proactive weed management strategies incorporating a mix of herbicides. Despite this, the overall environmental and economic gain from the use of GM crops has been, and continues to be, substantial.

Overall, there is a considerable body of evidence, in the peer-reviewed literature, and summarized in this chapter, that quantifies the positive economic and environmental impacts of crop biotechnology. The analysis in this chapter therefore provides insights into the reasons why so many farmers around the world have adopted and continue to use the technology. Readers are encouraged to read the peer-reviewed papers cited, and the many others who have published on this subject (and listed in the references section of the two main papers from Brookes and Barfoot that provided the background information for this chapter) and to draw their own conclusions.

LIFE BOX 1.1. NORMAN E. BORLAUG

Norman E. Borlaug (1914–2009) Nobel Laureate, Nobel Peace Prize, 1970; Recipient of the Congressional Gold Medal, 2007.



Norman Borlaug. Courtesy of Norman Borlaug.

The following text is excerpted from the book by biographer Leon Hesser, The Man Who Fed the World: Nobel Peace Prize Laureate Norman Borlaug and His Battle to End World Hunger, Durban House Dallas, Texas (2006):

From the day he was born in 1914, Norman Borlaug has been an enigma. How could a child of the Iowa prairie, who attended a one-teacher, one-room school; who flunked the university entrance exam; and whose highest ambition was to be a high school science teacher and athletic coach, ultimately achieve the distinction as one of the hundred most influential persons of the twentieth century? And receive the Nobel Peace Prize for averting hunger and famine? And could he eventually be hailed as the man who saved hundreds of millions of lives from starvation—more than any other person in history?

Borlaug, ultimately admitted to the University of Minnesota, met Margaret Gibson, his wife to be, and earned B.S., M.S., and Ph.D. degrees. The latter two degrees were in plant pathology and genetics under Professor E. C. Stakman, who did pioneering research on the plant disease rust, a parasitic fungus that feeds on phytonutrients in wheat, oats, and barley. Following 3 years with DuPont, Borlaug went to Mexico in 1944 as a member of a Rockefeller Foundation team to help increase food production in that hungry nation where rust diseases had taken their toll on wheat yields.

Dr. Borlaug initiated three innovations that greatly increased Mexico's wheat yields. First, he and his Mexican technicians crossed thousands of varieties to find a select few that were resistant to rust disease. Next, he carried out a "shuttle breeding" program to cut in half the time it took to do the breeding work. He harvested seed from a summer crop that was grown in the high altitudes near Mexico City, flew to Obregon to plant the seed for a winter crop at sea level. Seed from that crop was flown back to near Mexico City and planted for a summer crop. Shuttle breeding not only worked against the advice of fellow scientists, but serendipitously the varieties were widely adapted globally because it had been grown at different altitudes and latitudes and during different day lengths.

But, there was a problem. With high levels of fertilizer in an attempt to increase yields, the plants grew tall and lodged. For his third innovation, then, Borlaug crossed his rust-resistant varieties with a short-strawed, heavy tillering Japanese variety. Serendipity squared. The resulting seeds were responsive to heavy applications of fertilizer without lodging. Yields were six to eight times higher than for traditional varieties in Mexico. It was these varieties, introduced in India and Pakistan in the mid-1960s, which stimulated the Green Revolution that took those countries from near-starvation to self-sufficiency. For this remarkable achievement, Dr. Borlaug was awarded the Nobel Peace Prize in 1970.

In 1986, Borlaug established the World Food Prize, which provides $250,000 each year to recognize individuals in the world who are deemed to have done the most to increase the quantity or quality of food for poorer people. A decade later, the World Food Prize Foundation added a Youth Institute as a means to get young people interested in the world food problem. High school students are invited to submit essays on the world food situation. Authors of the 75 best papers are invited to read them at the World Food Prize Symposium in Des Moines in mid-October each year. From among these, a dozen are

sent for 8 weeks to intern at agricultural research stations in foreign countries. By the summer of 2007, approximately 100 Youth Institute interns had returned enthusiastically from those experiences, and all are on track to become productively involved. This is an answer to Norman Borlaug's dream.

Borlaug has continually advocated increasing crop yields as a means to curb deforestation. In addition to his being recognized as having saved millions of people from starvation, it could be said that he has saved more habitat than any other person.

When Borlaug was born in 1914, the world's population was 1.6 billion. During his lifetime, population has increased four times, to 6.5 billion. Borlaug is often asked, "How many more people can the Earth feed?" His usual response: "I think the Earth can feed 10 billion people, IF, and this is a big IF, we can continue to use chemical fertilizer and there is public support for the relatively new genetic engineering research in addition to conventional research."

To those who advocate only organic fertilizer, he says, "For God's sake, let's use all the organic materials we can muster, but don't tell the world that we can produce enough food for 6.5 billion people with organic fertilizer alone. I figure we could produce enough food for only 4 billion with organics alone."

One of Borlaug's dreams, through genetic engineering, is to transfer the rice plant's resistance to rust diseases to wheat, barley, and oats. He is deeply concerned about a recent outbreak of rust disease in sub-Saharan Africa which, if it gets loose, can devastate wheat yields in much of the world.

As President of the Sasakawa Africa Association (SAA) since 1986, Borlaug has demonstrated how to increase yields of wheat, rice, and corn in sub-Saharan Africa. To focus on food, population and agricultural policy, Jimmy Carter initiated Sasakawa-Global 2000, a joint venture between the SAA and the Carter Center's Global 2000 program.

Norman Borlaug has been awarded more than 50 honorary doctorates from institutions in 18 countries. Among his numerous other awards are the U.S. Presidential Medal of Freedom (1977); the Rotary International Award (2002); the National Medal of Science (2004); the Charles A. Black Award for contributions to public policy and the public understanding of science (2005); the Congressional Gold Medal (2006); and the Padma Vibhushan, the Government of India's second highest civilian award (2006).

The Borlaug family includes son William, daughter Jeanie, five grandchildren, and four great grandchildren. Margaret Gibson Borlaug, who had been blind in recent years, died on March 8, 2007 at age 95.

LIFE BOX 1.2. MARY-DELL CHILTON

Mary-Dell Chilton, Scientific and Technical Principal Fellow, Syngenta Biotechnology, Inc.; World Food Prize Laureate (2013); Winner of the Rank Prize for Nutrition (1987), and the Benjamin Franklin Medal in Life Sciences (2001); Member, National Academy of Sciences.

I entered the University of Illinois in the fall of 1956, the autumn that Sputnik flew over. My major was called the "Chemistry Curriculum," and was heavy on science and light on liberal arts. When I entered graduate school in 1960 as an organic chemistry major, still at the University of Illinois, I took a minor in microbiology (we were required to minor in something…). To my astonishment, I found a new love: in a course called "The Chemical Basis of Biological Specificity" I learned about the DNA double helix, the genetic code, bacterial genetics, mutations, and bacterial transformation. I was hooked! I found that I could stay in the chemistry department (where I had passed prelims, a grueling oral exam) and work on DNA under guidance of a new thesis advisor, Ben Hall, a



Mary-Dell Chilton in the Washington University (St. Louis) greenhouse in 1982 with tobacco, the white rat of the plant kingdom. Courtesy of Mary-Dell Chilton.

professor in physical chemistry. When Hall took a new position in the Department of Genetics at the University of Washington, I followed him. This led to a new and fascinating dimension to my education. My thesis was on transformation of *Bacillus subtilis* by single-stranded DNA.

As a postdoctoral fellow with Dr. Brian McCarthy in the microbiology department at the University of Washington, I did further work on DNA of bacteria, mouse, and finally maize. I became proficient in all of the then-current DNA technology. During this time, I married natural products chemist Prof. Scott Chilton, and we had two sons to whom I was devoted. But that was not enough. It was time to start my career!

Two professors (Gene Nester in microbiology and Milt Gordon in biochemistry) and I (initially as an hourly employee) launched a collaborative project on *Agrobacterium tumefaciens* and how it causes the plant cancer "crown gall." In hindsight, it was no accident that we three represented at least three formal disciplines (maybe four or five, if you count my checkered career). Crown gall biology would involve us in plants, microbes, biochemistry, genetics, protein chemistry, natural products chemistry (in collaboration with Scott), and plant tissue culture. The multifaceted nature of the problem bound us together.

My first task was to write a research grant application to raise funds for my own salary. My DNA hybridization proposal was funded. Grant money flowed in the wake of Sputnik. Our primary objective was to determine whether DNA transfer from the bacterium to the plant cancer cells was indeed the basis of the disease, as some believed and others disputed. We disputed this continually amongst ourselves, often switching sides! This was the start of a study that has extended over my entire career. While we hunted for bacterial DNA, competitors in Belgium discovered that virulent strains of *Agrobacterium* contained enormous plasmids (circular DNA molecules) which we now know as Ti (tumor-inducing) plasmids. Redirecting our analysis, we found that gall cells contained not the whole Ti plasmid but a sector of it large enough to encompass 10–20 genes.

Further studies in several laboratories worldwide showed that this transferred DNA, T-DNA, turned out to be in the nuclei of the plant cells, attached to the plant's own chromosomal DNA. It was behaving as if it were plant genes, encoding messenger RNA and proteins in the plant. Some proteins brought about the synthesis of plant growth hormones that made the plant gall grow. Others caused the plant to synthesize, from simple amino acids and sugars or keto acids, derivatives called "opines," some of which acted as bacterial hormones, inducing conjugation of the plasmid from one *Agrobacterium* to another. The bacteria could live on these opines, too, a feat not shared by most other bacteria. Thus, a wonderfully satisfying biological picture emerged. We could envision *Agrobacterium* as a microscopic genetic engineer, cultivating plant cells for their own benefit.

At that time, only a dreamer could imagine the possibility of exploiting *Agrobacterium* to put genes of our choice into plant cells for crop improvement. There were many obstacles to overcome. We had to learn how to manipulate genes on the Ti plasmid, how to remove the bad ones that caused the plant cells to be

tumorous, and how to introduce new genes. We had to learn what defined T-DNA on the plasmid. It turned out that *Agrobacterium* determined what part of the plasmid to transfer by recognizing a 25 base pair repeated sequence on each end. One by one, as a result of research by several groups around the world, the problems were solved. The Miami Winter Symposium in January 1983 marked the beginning of an era. Presentations by Belgian, German and two US groups, including mine at Washington University in St. Louis, showed that each of the steps in genetic engineering was in place, at least for (dicotyledonous) tobacco and petunia plants. Solutions were primitive by today's standards; but, in principle, it was clear that genetic engineering was feasible; *Agrobacterium* could be used to transform a number of dicots.

I saw that industry would be a better setting than my university lab for the next step: harnessing the Ti plasmid for crop improvement. When a Swiss multinational company, CIBA–Geigy, offered me the task of developing from scratch an agricultural biotechnology lab to be located in North Carolina where I had grown up, it seemed tailor made for me. I joined this company in 1983. CIBA–Geigy and I soon found that we had an important incompatibility: while I was good at engineering genes into (dicotyledonous) tobacco plants, the company's main seed business was (monocotyledonous) hybrid corn seed. Nobody knew whether *Agrobacterium* could transfer T-DNA. This problem was solved and maize is now transformable by either *Agrobacterium* or the "gene gun" technique. Our company was first to the market with Bt maize.

The company underwent mergers and spinoffs, arriving at the new name of Syngenta a few years ago. My role also evolved. After 10 years of administration, I was allowed to leave my desk and go back to the bench. I began working on "gene targeting," which means finding a way to get T-DNA inserts to go where we want them in the plant chromosomal DNA, rather than random positions it goes of its own accord.

Transgenic crops now cover a significant fraction of the acreage of soybeans and corn. In addition, transgenic plants serve as a research tool in plant biology. *Agrobacterium* has already served us well, both in agriculture and in basic science. New developments in DNA sequencing and genomics will surely lead to further exploitation of transgenic technology for the foreseeable future.

LIFE BOX 1.3. ROBERT T. FRALEY

Robert Fraley, Chief Technology Officer, Monsanto Co.; World Food Prize Laureate (2013); National Academy of Science Award for the Industrial Application of Science (2008); National Medal of Technology from President Clinton (1998).



Robert T. Fraley. Courtesy of Robert T. Fraley.

When I think back to my childhood on our family farm in central Illinois, I remember bailing hay and walking soybean fields to pull weeds. These pretty common farm jobs provided me with a perspective and the motivation to find better solutions to help farmers, like my dad, fight their most difficult problems. I am particularly grateful for my experience on our family farm because I learned firsthand both how challenging farming really is and how farmers continually adopt new and improved innovations.

It's humbling to remember life as a young farm boy, and then look at my career which progressed to pioneering research on gene transfer in plants and the development of Roundup Ready® soybeans and other biotech innovations. Although, from a very young age I knew I wanted to pursue a career in research, I had no idea then where science and innovation would take me, allowing me to travel the globe, interact with so many interesting people, in a career I truly enjoy.

Growing up in a rural setting, I attended a very small high school. In fact, I was the only student in my high school biology and physics classes. While a bit intimidated by how much one-on-one time I had with my science teacher, for me, this was an opportunity to grow, ask questions and absorb a new world of science and information. After graduating from high school, I received my Bachelor in Science at the University of Illinois, which established a sound foundation for my future. I continued my education at the U of I where I earned my Ph.D. in microbiology and biochemistry. I then spent 2 years of postdoctoral fellowship research at University of California, San Francisco, where I studied ways to introduce genes into plant and animal cells using liposomes. This was the period where I became focused on how biotechnology could be used to improve crops.

In 1981, Dr. Ernie Jaworski hired me to join a small, but talented team of scientists at Monsanto. It was exciting to work with this team. I valued our collaborative efforts to address some of agriculture's greatest challenges. Ironically though, our research started by using *Agrobacterium* to introduce new genes into the petunia, not your traditional crop! Looking back, this was a great decision because we were able to quickly prove the science. The petunia became the first genetically engineered plant, and it laid the foundation for many innovations in agriculture, including plants with resistance to pests, increased crop yields and protection against drought, and other environmental conditions.

As we advanced the research and technology, we developed solutions to help farmers address challenges on their farms. We shared our research results and safety analyses with the scientific community, regulatory bodies around the world and our farmer customers. Excitement supporting the science continued to spread and our team became recognized as key contributors to the worldwide scientific and agriculture communities. This was very humbling and led to an experience I will never forget, receiving the National Medal of Technology from President Clinton in 1998.

Looking back on all this though, we didn't do a great job of communicating directly with consumers and because of that, years later, we continue to work to address common misperceptions about how food is grown and if it is safe, nutritious, and sustainable. As a scientist, I was comfortable letting the evidence speak for itself. Although not joining the conversation with consumers earlier is my greatest regret, I am pleased that we have since engaged in this dialogue and continue to find common ground.

Throughout my career at Monsanto, I've held several roles within the Technology organization. My current role as Chief Technology Officer continues to excite me because I am not only leading a team of the top scientists in the ag industry, but I have the privilege of talking with farmers and seeing the process from beginning (in the lab) to end (on our customers' fields). One opportunity that has been especially rewarding for me in the last couple of years is engaging with broad audiences and furthering the dialogue with consumers, as well as partners like the Gates Foundation, Clinton Global Initiative and Conservation International. I see the opportunity to join the conversation with new and diverse groups as an important step in the right direction.

As I look back, the recognition that means the most to me as a scientist is the World Food Prize. The acknowledgment that biotechnology has made an important contribution to world food security was very rewarding and my close relationship with Dr. Norman Borlaug made this award even more special and personal. I have always admired Dr. Borlaug and the impact his scientific leadership provided. He emphasized the significance of food security and always impressed on me the need to think globally and forward for future generations. This is particularly critical today, as we face one of mankind's greatest challenges. By 2050, our global population will swell to 9.5 billion people, so we will need to produce more food in the next 30–35 years than we have in the entire history of the world. Dr. Borlaug said, "Food is the

moral right of all who are born into this world," and by using agriculture effectively, we can address poverty, hunger, and overcome some of our biggest obstacles. Dr. Borlaug's leadership and mentorship continue to have a great impact on me, my career, and my world view.

The agriculture industry holds great growth potential and is at the center of so many of today's challenges—mitigating climate change, environmental impacts, growing population, changing diets, and food production demand. Continued innovation, both in biology and data science, can transform agriculture globally. I believe that we can not only meet the challenge of food security but also sustainably increase production to the point where we can reduce farming's footprint around the world. It is a very exciting time to be involved in agriculture and I encourage all who are interested in science to consider career opportunities in this industry. The innovation and developments that create sustainable solutions for farmers can lead to fulfilling and rewarding careers.

Roundup Ready is a registered trademark of Monsanto Technology, LLC. © 2015 Monsanto Company.

REFERENCES

American Soybean Association (2001): *Conservation Tillage Study*. Available at: http://www.soygrowers.com/ctstudy/ctstudy_files/frame.htm (accessed on September 8, 2015).

Brookes G, Barfoot P (2014a): Economic impact of GM crops: the global income and production effects 1996–2012. *GM Crops Food* **5**: 1, 65–75. Available at: www.landesbioscience.com (accessed on September 8, 2015).

Brookes G, Barfoot P (2014b): Key global environmental impacts of GM crop use 1996–2012. *GM Crops Food* **5**: 2, 1–12. Available at: www.landesbioscience.com (accessed on September 8, 2015).

Conservation Tillage and Plant Biotechnology (CTIC) (2002): *How New Technologies Can Improve the Environment by Reducing the Need to Plough*. Available at: http://www.ctic.purdue.edu/CTIC/Biotech.html (accessed on September 8, 2015).

James C (2012): *Global Status of Transgenic Crops: 2013, Brief Number 46*. International Service for the Acquisition of Agri-Biotech Applications (ISAAA). Available at: http://www.isaaa.org/resources/publications/briefs/46/default.asp (accessed on September 8, 2015).

Jasa P (2002): *Conservation Tillage Systems*. University of Nebraska. Available at: agecon.okstate.edu/isct/labranza/jasa/tillagesys.doc (accessed on September 8, 2015).

Kovach J, Petzoldt C, Degni J, Tette J (1992): *A Method to Measure the Environmental Impact of Pesticides*. New York's Food and Life Sciences Bulletin. NYS Agricultural Experiment Station, Cornell University, Geneva, NY, p 139, 8 pp annually updated. Available at: http://www.nysipm.cornell.edu/publications/EIQ.html (accessed on September 8, 2015).

Lazarus WF (2012): *Machinery Cost Estimates May 2012*. University of Minnesota Extension Service. Available at:http://www.minnesotafarmguide.com/news/regional/machinery-cost-estimates/pdf_a5a9623c-636a-11e3-8546-0019bb2963f4.html (accessed on September 8, 2015).

Reeder R (2010): *No-Till Benefits Add Up with Diesel Fuel Savings*. Available at: http://www.thelandonline.com/currentedition/x1897235554/No-till-benefits-add-up-with-diesel-fuel-savings (accessed on September 8, 2015).

University of Illinois (2006): *Costs and Fuel Use For Alternative Tillage Systems*. Available at: www.farmdoc.uiuc.edu/manage/newsletters/fefo06 07/fefo06 07.html (accessed on September 8, 2015).

USDA (2013a): *An Online Tool for Estimating Carbon Storage in Agroforestry Practices (COMET-VR)*. Available at: http://www.cometvr.colostate.edu/ (accessed on September 8, 2015).

USDA (2013b): *Energy Estimator: Tillage*. Available at: http://ecat.sc.egov.usda.gov (accessed on September 8, 2015).

CHAPTER 2

Mendelian Genetics and Plant Reproduction

MATTHEW D. HALFHILL

Department of Biology, Saint Ambrose University, Davenport, Iowa

SUZANNE I. WARWICK

Agriculture and Agri-Food Canada, Eastern Cereal and Oilseeds Research Centre, Ottawa, Canada

2.0. CHAPTER SUMMARY AND OBJECTIVES

2.0.1. Summary

Plant biotechnology is an incredibly dynamic and modern field. Future innovations will come from genomics, bioinformatics, epigenetics, and emerging technologies that are not yet imagined. It is easy to be charmed by the modernity of the field as we generate novel plant material never before seen in nature. This chapter is designed to remind us that after we genetically modify a location in the genome with a novel deoxyribonucleic acid (DNA) sequence, that innovation will conform to ancient genetic processes illuminated by Gregor Mendel over 150 years ago. Every new trait that plant biotechnologists coax into a genome will become just another gene of the thousands of genes already present in the organism and will be passed to the next generation following the rules of plant reproduction. Effective plant biotechnologists have a solid foundation in basic genetics and plant reproductive biology, and they utilize this foundation to change the organism through biotechnology for scientific advancement.

2.0.2. Discussion Questions

1. What is a gene, and why are there multiple viable definitions?
2. How does the discrete nature of chromosomes impact sexual reproduction in plants?
3. What would be the consequence of sexual reproduction if mitosis was the only form of cell division?
4. How do the reproductive features of plants regulate the degree of inbreeding?

2.1. OVERVIEW OF GENETICS

Flowering plants (angiosperms) and conifers (gymnosperms) are diverse organisms that have conquered the terrestrial world, making the planet green. Angiosperms comprise the most important crop and horticultural plants, while gymnosperms are important in forestry. Gregor Mendel, the

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

nineteenth-century monk, was the first person to demonstrate the inheritance of genes (even though he did not know what genes were in the molecular sense) using the garden pea plant. His research is the basis of inheritance theory and practice. Plants have sundry methods of reproduction ranging from vegetative propagation to sex by cross-fertilization, which sets them apart from the relatively mundane world of animal reproduction. With the incredible diversity of reproduction methods, plants maintain genetic variation in various ways.

The field of genetics has impacts on all aspects of the science of biology, but individual disciplines within biology utilize different types of genetic information. In order to discuss plant reproduction specifically, several universal genetic definitions must be introduced. In its simplest definition, the field of genetics is the study of genes. DNA is the genetic material in organisms that stores all the information that encodes for life. The sequence of nucleotides (DNA building blocks: A (adenine), C (cytosine), G (guanine), and T (thymine)) stores the instructions to produce proteins and information that allow for the regulation of the genetic material. The DNA sequence serves as a type of software or programming language that the cell uses to produce and regulate all the necessary products for life. DNA exists as a double helix, and each nucleotide is paired with its complementary base making a base pair (adenine with thymine, cytosine with guanine).

For this chapter, a gene is defined as a contiguous sequence of DNA that contains regulatory regions and a sequence that most often encodes for a protein. Many sequences in the genome of an organism are outside this definition of the gene, and in fact, much of a plant's DNA would not be considered to be part of any "gene" as defined earlier. The next level of genetic organization is the chromosome, which is an organization unit for a single molecule of DNA and associated proteins that reside within the nucleus. The chromosome-associated proteins help package and condense DNA for packing into the nucleus of a cell. The genome of an organism is the entire sequence of DNA inclusive of all the chromosomes. DNA is also present within certain cellular organelles: the mitochondria and chloroplasts. Plants therefore contain three distinct genomes—the nuclear, mitochondrial, and chloroplast, and this chapter will focus specifically on the DNA contained within the nucleus. If we draw an analogy comparing genetics to the structure of this book, nucleotides are similar to letters that form three letter words. Genes are similar to sentences, and chromosomes are similar to chapters. The genome is similar to a complete book, and a library would be a collection of different species (see Chapter 6 for detailed explanation on molecular genetics).

Molecular, cellular, organismal, population, and evolutionary studies all have genetic components, and build on traditional knowledge about genes. For molecular research, the DNA sequence of a gene and its presence and role within the genome are critically important. The sequence itself determines how a gene functions and impacts on the final characteristics of the organism. In larger scale research, such as population and evolutionary studies, both the transcribed DNA within a gene and that which falls outside genes (spacer regions) may be used to describe population structure. Often in a comparative study, the sequences within the genes are highly conserved that is too similar in makeup, and are therefore non-informative with respect to deciphering genetic relatedness. With this in mind, variable genetic information outside the genes is often more useful for large-scale population studies. These DNA sequences are often used in various types of DNA fingerprinting procedures to elucidate differences between populations. It should be noted that there are differences of opinion on basic definitions of critical terms such as "gene." Unlike our definition, some scientists/researchers refer to the gene as simply the coding region (without the DNA responsible for regulating gene expression). Others have a broader view of the gene to encompass nearly any stretch of DNA. Genetics is a dynamic field whose terminology can be confusing—almost like a rapidly evolving language.

For plant reproduction, the most important genetic level is the chromosome, since chromosomes are the largest units of DNA passed from parents to offspring (progeny). In other words, this chapter is the story of chromosomes. In plants as in all eukaryotes (organisms with a nucleus), chromosomes are linear pieces of DNA that have a single centromere and two arms (Fig. 2.1). The centromere is the constricted region of the chromosome and serves as a connection between the chromosome arms.

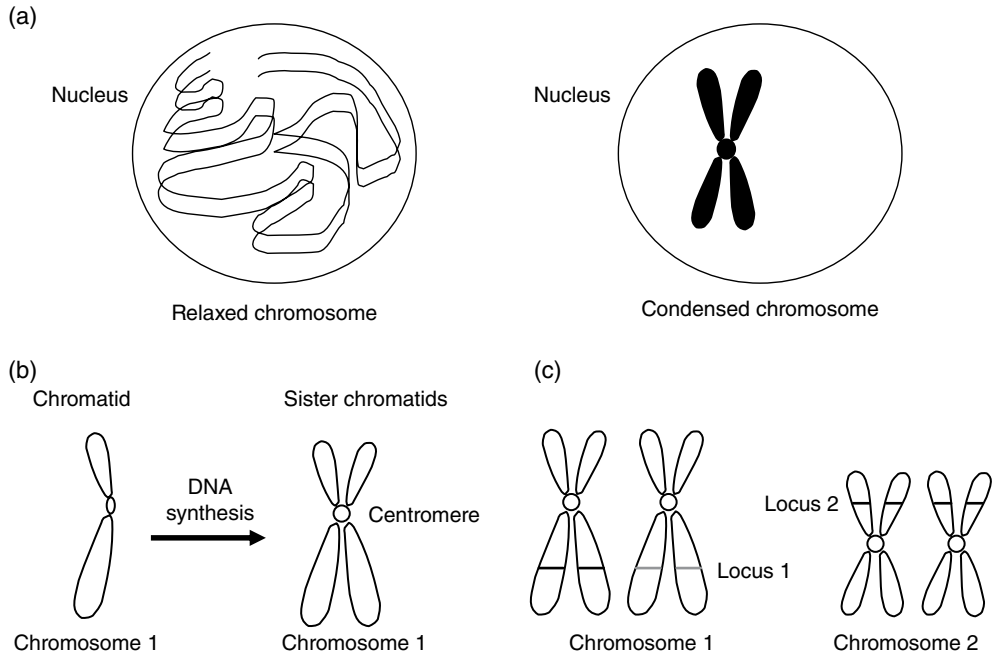


Figure 2.1. Chromosomes have several physical states during the life of a cell: (a) chromosome physical states, (b) chromosome conformations, and (c) homologous chromosomes.

Centromeres also play an important role in cell division, which will be discussed later in the chapter. The genes mainly exist on the chromosome arms. Different plant species vary widely in chromosome number, and this number often defines a species as being different from another. The number of chromosomes within a nucleus is defined as the ploidy of the cell. For example, the model plant *Arabidopsis thaliana* has a total of 10 chromosomes (5 pairs), while the crop plant soybean (*Glycine max*) has 40 chromosomes (20 pairs). Some plants have tremendously large genomes. For example, some lilies have hundreds of chromosomes. Chromosomes vary in length, that is, in the number of nucleotides that make up the DNA molecule, and therefore vary in size when visualized under the microscope. Each chromosome has hundreds to thousands of genes contained within the sequence of DNA, along with sequences between the genes. This connecting DNA has been historically called "junk" DNA, but current research is discovering that intergenic DNA sequences may play several critical roles such as regulating how genes and chromosomes interact at higher levels.

To understand biotechnology and genetics, it is essential to define and understand how chromosomes exist within the nucleus. Chromosomes are organized in two different basic physical structures during the life of the cell. During most of the cells' adult life, the chromosome exists in a relaxed state, where the DNA is loosely wrapped around chromosomal proteins (Fig. 2.1a). This physical state allows the DNA to be read (transcribed and translated) so that the appropriate proteins are produced. As the chromosomes prepare for cell division, they become tightly wound around chromosomal proteins and are described as being in the condensed state (Fig. 2.1a). Chromosomes can only be visualized under the light microscope when they are condensed. During different points in the cell cycle, chromosomes may be in different conformations. Initially after cell division, a chromosome exists as a single molecule of double-stranded DNA with a single centromere, which is called a chromatid (Fig. 2.1b). After the DNA synthesis phase of the cell cycle, the chromosome exists as two molecules of identical double-stranded DNA connected at the single centromere. The two DNA molecules within a chromosome are called sister chromatids, and they stay connected

until they are separated by one of the types of cell division. DNA synthesis does not represent a change in the total chromosome number, as chromosome numbers remain the same during the lifetime of the plant. A single chromosome then may exist in either a pre-replicated (one chromatid) or replicated state (two sister chromatids). The different states of chromosomal arrangements within the life of a cell will be important as we describe cell division and sexual reproduction.

Most cells in a plant have two copies of each chromosome, which are called homologous chromosomes or a chromosome pair (Fig. 2.1c). Generally speaking, one of the individual chromosomes in a pair is derived from the maternal parent and one from the paternal parent. Gender identity and parenting is sometimes confusing to think about in plants that have the ability to self-fertilize (when the same plant's pollen fertilizes the ovum), but one of the homologous chromosomes comes from the pollen and one from the ovum even if all the chromosomes come from the same plant. Hermaphrodites (organisms with both male and female organs) and selfing are considered to be anomalies in the animal kingdom but are frequent among plants. As we will discuss later in this chapter, plants have a wide array of reproductive strategies to achieve the pairing of the chromosomes.

Most adult plant cells have two copies of all chromosomes, and the ploidy level is defined as the diploid state (2N). In order to sexually reproduce, the total chromosome number is divided in half, and this reduced chromosome number in the sexual gametes is defined as the haploid state (N). During most of an angiosperm plant's life, the diploid sporophyte stage dominates and produces diploid cells during cell division. In the small reproductive structures (pollen grains and ovaries), the haploid gametophyte stage is present and gives rise to haploid sex cells. Even with the diversity of chromosome numbers observed among plant species, eukaryotic chromosomes function under the same rules during cell division. During normal cell division (mitosis) in the sporophyte, the chromosome number is maintained in the diploid state. During the production of the gametophyte (meiosis), the two copies of each chromosome separate from one another and produce cells with half the normal number of chromosomes. All the variations of reproductive mode are simply complexities of how the two homologous chromosomes come together during the process of reproduction.

2.2. MENDELIAN GENETICS

Gregor Mendel, a member of the Augustinian monastery in Brno, Moravia (within the current Czech Republic borders), was the first person to describe how chromosomes are transmitted between generations (Fig. 2.2). Mendel combined what are now considered typical plant-breeding procedures, such as keeping accurate records of the characteristics that appeared in the offspring of selected parents and the control of pollination of the experimental plants, with statistics to describe how traits behave over generations. The molecular basis of genetics was not understood in the 1800s, but Mendel observed and recorded the phenotypic traits within the plants that he grew on the grounds of the monastery. The phenotype is the physical appearance of an organism, and the genotype is the underlying genetic makeup of an organism. Using pea plants (*Pisum sativum*), Mendel was able to track the segregation of traits over generations, and thus indirectly described the laws of how chromosomes act within cells. He accurately described the cellular process of chromosomal segregation without the benefit of knowing what was occurring within the nucleus or that chromosomes existed. Gregor Mendel's work in genetics was relatively obscure in his own day but was "rediscovered" in the twentieth century.

Mendel made a good choice when he decided to work with peas, since the pea plants he used differed from one another in several relatively simple phenotypic traits. Seed shape and color, pod shape and color, plant height, and flower position were the traits that he traced over generations of sexual reproduction (Mendel 1866). The pea plants had different variants for a given trait (Fig. 2.3). For example, some of the pea plants had yellow seeds, while others had green seeds. Each of the traits that Mendel followed was controlled by a single gene and the traits themselves were often discrete.



Figure 2.2. Gregor Mendel was the father of genetics. (*Source*: Iltis, https://commons.wikimedia.org/wiki/File:Gregor_Mendel_oval.jpg. Used under CC-BY 4.0 https://creativecommons.org/licenses/by/4.0/deed.en)

Trait	Dominant	Recessive
Seed color	Yellow - *Y*	Green - *y*
Seed form	Round - *R*	Wrinkled - *r*
Seed coat color	Grey - *A*	White - *a*
Pod form	Inflated - *V*	Restricted - *v*
Pod color	Green - *G*	Yellow - *g*
Flower position	Axial along stem - *F*	Terminal on top - *f*
Stem length	Tall - *L*	Short - *l*

Figure 2.3. Traits of the pea plant used by Mendel to discover the genetic laws of segregation and independent assortment. Each trait had two phenotypes: one controlled by a dominant allele and the other controlled by a recessive allele.

That is, seeds could be scored as either yellow or green, and not a mixed or splotched variant that was in between the original parents.

Mendelian traits are controlled by a single gene, and therefore the protein product from a single gene directly leads to the characteristic phenotype. Mendelian traits may have multiple different versions that make different proteins with varying characteristics, but the gene that controls the trait is at a single location within a chromosome in the genome called a locus (Fig. 2.1c). The different versions of each gene are called alleles, and they differ from one another in the sequence of DNA at that chromosomal locus. Mendelian traits are also defined by having discrete variation, where the different phenotypes of the trait can be broken into obvious categories. In the example of pea plant height, tall versus short plant type is determined by the genotype at a single genetic locus that controls height.

As you will see throughout this book, most traits are more complex than Mendelian traits because they are controlled by the gene products of many genes, and hence called polygenic traits. Polygenic traits exhibit continuous variation, where the trait can show a wide range of phenotypes. Multifactorial traits are controlled by multiple genes (polygenic traits) and the environment in which the plant is grown. Multifactorial traits also exhibit continuous variation, and will vary based on the environmental conditions. Polygenic and multifactorial traits will be discussed specifically in the breeding chapter of this book. The traits that Mendel followed had two specific characteristics; they had discrete variation and were controlled by the action of a single gene.

Mendel was very observant, and understood his plants. Another reason his choice in using peas was good is that peas normally self-fertilize, which made all of his interpretations of transmission genetics much simpler than if he'd picked a species that was normally (or even partially) outcrossers. He used plant lines that would only generate plants of a single type when the plants were allowed to self-fertilize. These plants were homozygous for that trait, which meant that the two homologous chromosomes had the same allele. When homozygous plants are selfed, the resulting progeny are always homozygous. Mendel's method to track segregation was based on crossing plants that were homozygous and differed for the phenotypic trait of interest. For example, he would cross (instead of selfing) plants that were homozygous yellow and homozygous green for seed color, and then record the phenotypic ratio in progeny of each subsequent generation.

By crossing different homozygotes, Mendel generated plants whose two homologous chromosomes each had a different allele of the gene (Fig. 2.4a). The condition of having two different alleles in a single gene is called heterozygous. All the plants generated from the initial cross (F_1 hybrids or F_1 generation) would have the same genotype, but could have either one of two different parental phenotypes. In the heterozygous plants, Mendel discovered that certain variants of a trait appeared to mask or cover the expression of other variants. A variant that would cover the other type was termed "dominant," while the phenotypic type that would disappear was called recessive. When we write allele names, we often use uppercase letters for dominant alleles and lowercase letters for recessive alleles. We understand today that dominant alleles have a sequence of DNA that encodes

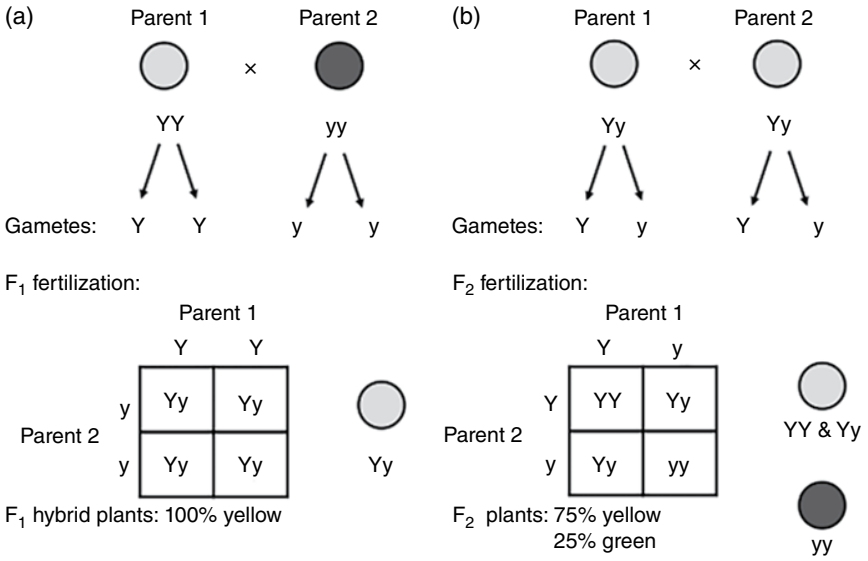


Figure 2.4. A monohybrid crossing system involving a single-gene model, in which the two alleles segregate from one another in the production of gametes: (a) monohybrid cross between a homozygote dominant plant and a homozygote recessive plant and (b) self-fertilization of a heterozygous F_1 hybrid. Results of crosses are represented by Punnett Squares.

for a functional protein, while many recessive alleles have changes in the DNA sequence, called mutations, which cause the encoded protein to be nonfunctional. Therefore, in a heterozygous plant, functional and nonfunctional proteins are produced, and the plant has the phenotype of the dominant allele resulting from the functional protein. In Mendel's experiments, he would see that the dominant trait would mask the expression of the recessive trait.

After crossing the homozygous parents and generating a heterozygous hybrid plant (F_1), Mendel would allow the hybrid plant to self-fertilize. In the subsequent F_2 plants or F_2 generation, plants with the recessive trait would reappear (Fig. 2.4b). Mendel realized that the recessive allele was not replaced or destroyed by the dominant allele, but its phenotype was just masked in the heterozygous individuals. With his intricate record keeping of counting the plants with different phenotypes, Mendel observed that the dominant plants occurred in 75% of individual F_2 plants, while recessive plants occurred at a frequency of 25%. Mendel's crosses may be visualized in a graphical table called a Punnett Square that depicts the number and variety of genetic combinations in a genetic cross (Fig. 2.4). The latter was named after Reginald Punnett, who worked with William Bateson to confirm experimentally the findings of Gregor Mendel. Their investigations of the exceptions to Mendel's rules led to the discovery of genetic linkage in the pea, which will be discussed later in this chapter. Using a Punnett Square, the possible genotypes of the gametes from each parent are placed on adjacent axes, and the matrix within the Punnett Square represents all possible outcomes from sexual reproduction.

Using his crossing data, Mendel realized that plants contained two copies of genetic material. Although he did not know that each plant had two different sequences of DNA on the two homologous chromosomes, he could predict the expected segregation frequencies over all the traits that he tracked over multiple generations. The fundamental process that Mendel discovered was that plants contained two versions of every gene, and that those genes were discrete particles that could separate from one another over the generations.

2.2.1. Law of Segregation

In his crosses using single traits, or monohybrid crosses, Mendel described the first of his genetic laws describing how traits are passed between generations. He did not know that DNA controlled the traits he observed, but we will state his law based on current knowledge that DNA is genetic material and is stored in chromosomes. Based on the fact that dominant and recessive alleles segregate from one another in progeny derived from heterozygous plants, he described the Law of Segregation, which states that two homologous chromosomes separate from one another during the production of sex cells. In practical terms, this means that half of the sex cells will be produced with one allele and half with the other allele in a heterozygous plant.

2.2.2. Law of Independent Assortment

Mendel also crossed plants that differed at multiple traits at the same time. When plants that differed at two traits were crossed, or dihybrid crosses, Mendel determined that the traits segregated independently from one another (Fig. 2.5). This phenomenon was described in the Law of Independent Assortment, where chromosomes from different homologous chromosome pairs separate independently from one another during the production of sex cells. Chromosomes are independent molecules of DNA, and only homologous chromosomes pair with one another during gamete production. Therefore, nonhomologous chromosomes will divide completely randomly into the daughter cells.

It is an interesting historical fact that the traits that Mendel studied were controlled by genes on different chromosomes (or far enough away from one another in the genome to act independently). This is often deemphasized when discussing Mendel's work and it should not be, because if the genes had been on the same chromosome, his results would have been different. Genetic linkage, or the fact that genes on the same chromosome tend to be inherited together, would have caused linked alleles and corresponding traits to remain together rather than segregate independently. He did not

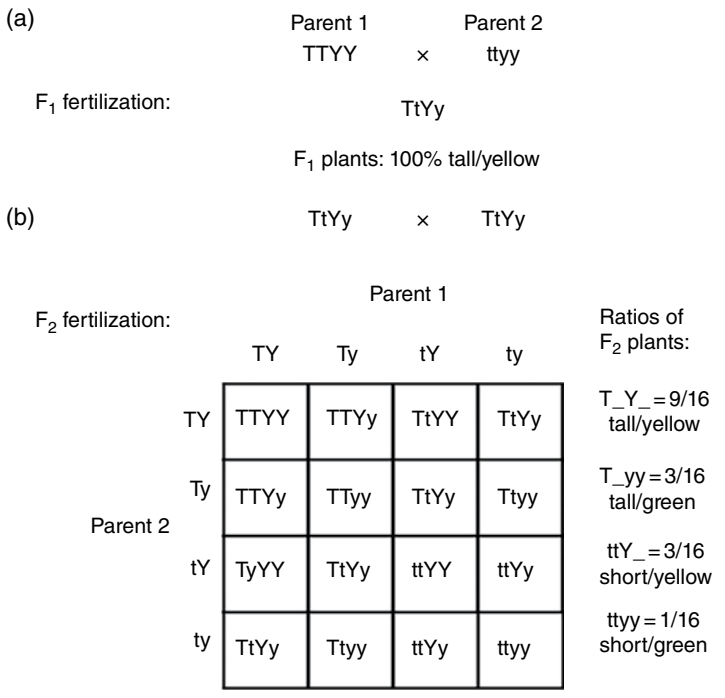


Figure 2.5. A dihybrid crossing system involving a two-gene model where the alleles of two genes independently assort from one another in the production of gametes: (a) dihybrid cross and (b) F_1 self-fertilization. Dihybrid crossing system involving a two-gene model where the alleles of two genes independently assort from one another in the production of gametes.

understand it at the time, but Mendel's traits were each controlled by a single gene on completely different chromosomes, which allowed them to segregate in the patterns he observed.

There were numerous experiments in the crossing of different species or varieties of plants during the eighteenth and nineteenth centuries with the primary intention to obtain new and improved varieties of fruits and vegetables. Knight (1799) and Goss (1824) in the United Kingdom both worked on pea, and, in fact, made the same types of crosses as Mendel. Each researcher observed the same general segregation patterns, but did not record the same numbers as did Mendel. Knight chose pea as a research plant because of its short generation time, the numerous varieties available, and the self-fertilizing habit, which made the protection of flowers from insects carrying pollen unnecessary. Presumably, Mendel had the same goals and rationale.

Mendel's laws have served as the basis for all fields of genetics. Of course, once the structure of DNA was described by Watson and Crick in 1953, the age of modern genetics began. Even though the mechanisms as to how DNA could store genetic information was not known, Mendel's principles still correctly described how genes were transferred between generations. Mendel's important work illustrates that comprehensive knowledge on a subject is not needed to make an important contribution in science. To continue our discussion of plant reproduction, we must describe the two types of cell division that separate chromosomes from one another during the life of the cell.

2.3. MITOSIS AND MEIOSIS

Mendel's observations and subsequent research prompted cell biologists to study the movement of chromosomes during the process of cell division. Plant growth and sex cell production are the result of two different types of cell divisions: chromosome copying (mitosis) and chromosome reducing

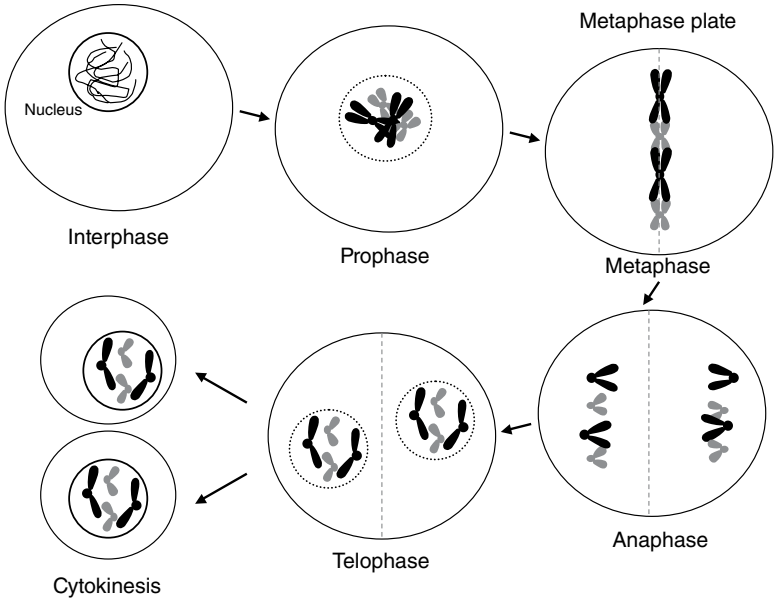


Figure 2.6. The stages of mitosis based on arrangement of the chromosomes.

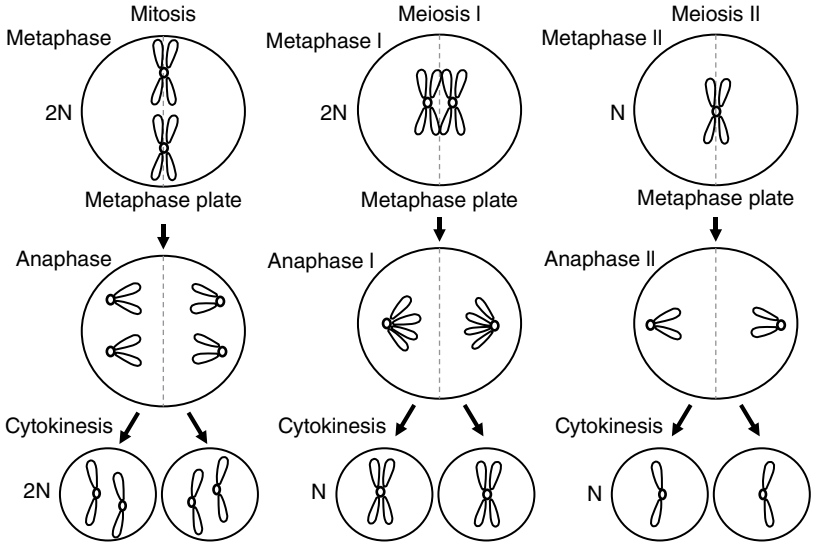


Figure 2.7. Mitosis and the two steps of meiosis differ from one another by the arrangement of the homologous chromosomes prior to cell division.

(meiosis). Most cells in a plant and any other complex organism go through an exact copying process in which the original chromosome number remains the same. This process that allows simple plant growth is called mitosis, in which a cell divides into two exact copies of the original (Fig. 2.6) with identical DNA sequences. In mitosis, the chromosome number is maintained in each daughter cell as a result of the division of sister chromatids at the centromere.

In order to proceed through sexual reproduction, cells must undergo the process of meiosis, a form of cell division where the resultant cells have half (haploid) the total number of chromosomes (Fig. 2.7) and with, therefore, different DNA sequences. If the chromosome number was not reduced in sex cells (gametes), the number of chromosomes would double after each generation of sexual reproduction. This of course is not the case, as each plant species generally retains its chromosome

number over generations. Meiosis allows for two haploid cells to join during fertilization to reconstitute the two copies of each chromosome in the progeny. Mitosis and meiosis are the two processes by which a cell may go through cell division, and each process has a different goal based on the total number of chromosomes required in the daughter cells.

2.3.1. Mitosis

The process of mitosis has a goal of maintaining the complete number of chromosomes and identical DNA sequences during cell division. Mitosis is a highly ordered process, because chromosome loss during cell division would be detrimental to the adult plant. Mitosis can be broken into five basic steps each defined by the organizational state of the chromosomes (Fig. 2.6).

The chromosomes are in the relaxed state throughout most of the life of the cell, which is called interphase. Interphase is the period of cellular life when the cell grows and prepares its chromosomes for cell division. During the synthesis phase (S phase) of interphase, chromosomal DNA is replicated to form sister chromatids. As the cell enters mitosis, the chromosomes condense into the tightly wound state and the nucleus breaks down, which are characteristics of prophase. The chromosomes appear in a disorganized mass that can be seen under the light microscope. The cellular machinery that performs the actual work of cell division involves a group of proteins called the mitotic spindle apparatus, but we will focus on the state of the chromosomes during mitosis in this chapter. As the chromosomes become organized along the middle of the cell, they enter metaphase. During metaphase, the chromosomes line up at the center of the cell with each of the sister chromatids on opposite side of the metaphase plate. The centromere sits directly on the middle line, and is broken in half and pulled to the opposite ends of the cell during anaphase. The chromosomes appear as small V's, with the centromere being pulled to the opposite poles with chromosome arms lagging behind. During this phase, the cell transiently has a $4N$ chromosomal number, because the centromeres between the sister chromatids are broken producing two chromosomes. When the chromosomes reach the opposite ends of the cell, the nuclear membranes re-form, which characterizes telophase. At this point, the two sister chromatids from all the chromosomes have been separated from one another, and the cell can divide by a process called cytokinesis into two daughter cells that have the exact same DNA. During mitosis, the chromosomes are broken at the centromere and the two daughter cells each acquire a complete copy of the cell's genome.

2.3.2. Meiosis

Meiosis is the type of cell division used to make sex cells or gametes. The goal of meiosis is to generate haploid cells, which have half the number of chromosomes as the original cell. Meiosis is a two-step process, where the original cell goes through two divisions in order to make haploid cells. In the first division (I), homologous chromosomes line up together and separate from one another to generate haploid cells. In the second meiotic division (II), sister chromatids of each chromosome divide in a process identical to mitosis. It can be said that meiosis simply adds a reductive division to separate the homologous chromosomes, and then goes through a mitotic division of remaining chromosomes.

The two meiotic divisions have similar steps to that already described in mitosis, with the condensation of the chromosomes, alignment in the center of the cell, being pulled to opposite poles, and then cell division. The differences occur in how the homologous chromosomes interact with one another (Fig. 2.7). In the first meiotic division, the homologous chromosomes find one another and form a structure called the tetrad. During prophase I, the homologous chromosomes interact with one another, which allows for the transfer of genetic material between the homologous chromosomes in a process known as crossing over or recombination. Recombination of this fashion generates diversity when the homologous chromosomes swap DNA. Metaphase I is also different in meiosis, as the homologous chromosomes in the tetrad straddle the metaphase plate, with each chromosome on one side. During anaphase I, complete homologous chromosomes, each with their two sister chromatids, are pulled to the opposite poles of the cell. The centromere remains completely intact as

each separate homologous chromosome is pulled to the opposite end of the cell. After cell division, each daughter cell has only one of each homologous chromosome, and therefore only half of the genetic material. The first meiotic division results in a reduction of genetic material by half.

The second meiotic division is exactly like mitosis, but with half the genetic material per cell, with the chromosomes lining up at the metaphase plate with the sister chromatids on each side of the cell. The centromeres are then broken, and the sister chromatids are pulled to opposite ends of the cell. This division results in two cells with identical genetic material, which is exactly the same process as mitosis, except with a haploid number of chromosomes. Meiosis and mitosis are similar processes but differ in how the chromosomes are pulled apart. In mitosis, the complete genome is retained in the daughter cells, while meiosis reduces the genome size in half by separating the homologous chromosomes. Therefore, growth is achieved by mitosis as numerous exact copies of the diploid cells are made, allowing for each cell to function in the adult plant. Meiosis prepares for sexual reproduction by generating haploid cells, which will be combined by the process of fertilization with other haploid cells to reconstitute the normal number of two homologous chromosomes.

2.3.3. Recombination

Recombination or the crossing over of DNA between chromosomes during meiosis is a critically important process that generates genetic diversity in plant species. If recombination did not occur, each chromosome would be essentially static and immortal with the same alleles always linked together on the same piece of DNA. The only changes that could occur in the DNA sequence would be caused by mutation, and each mutation would stay on the same piece of DNA forever. If this were the case, then plant improvement via breeding would be impossible. In both nature and agriculture, the "goal" is to combine advantageous alleles together within the same breeding line to improve a plant for natural or agricultural settings. Without recombination, the target of selection would be the chromosome with the allele of interest, and there would be a limited number of chromosome combinations from which to make selections. Luckily for crop breeders, mutation is not the only process that generates genetic diversity.

Recombination allows for alleles to be shuffled during every meiotic division (Fig. 2.8). It has been estimated that crossing over occurs during every meiotic division for each chromosome, and therefore the life span of any chromosomal sequence is actually only one generation. This allows for different alleles at different chromosomal loci to reshuffle and land on the same chromosome. Crop breeders rely on this process, because they attempt to select for recombination events that liberate the specific allele from a genetic background to improve the crop line without having to select for chromosomes. Oftentimes, crop plants have been highly selected to have a group of alleles that help the crop to perform well under specific agricultural conditions. A single new allele may make the crop better, but the breeder needs to retain all the original genes of the crop. The process of recombination allows the

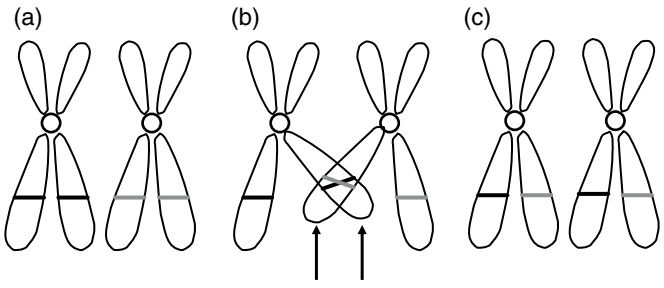


Figure 2.8. Recombination occurs when homologous chromosomes trade DNA sequences, thus generating genetic diversity: (a) before crossing over, (b) during crossing over, and (c) after crossing over.

breeder to try to find specific recombination events where the one allele has crossed over to join all the other original crop-selected alleles (see Chapter 3 for an in-depth description of plant breeding).

2.3.4. Cytogenetic Analysis

Scientific methods to observe chromosomes have improved greatly since Mendel outlined the laws that describe chromosome movement across generations. The easiest way to observe chromosomes is via chromosome staining during mitosis. Many readers can remember back to their high school biology classes where they observed stained onion (*Allium cepa*) root tips with the microscope. In these lab exercises, condensed chromosomes were stained with a DNA specific dye (a fuchsin-based DNA-specific stain developed by Feulgen in 1914), and the different stages of cellular mitotic division determined by observing the patterns of the chromosomes in each cell. Chromosome viewing by simple light microscopy is, however, limited to those plant species with large chromosomes in which single layers of actively dividing cells can easily be attained. These conditions are not common to most tissue types in adult plants.

More advanced cytogenetic techniques to observe chromosomes have been developed over the last 50 years, and are now being combined with molecular tools in the field of plant genomics research. Fluorescence *in situ* hybridization (FISH) is a method that utilizes small fluorescently labeled DNA fragments to paint different chromosomes (Fig. 2.9). In this technique, nuclear DNA is fixed to the surface of a slide preparation and the labeled DNA fragments bind to chromosomes with homologous complementary sequences. Since the chromosomes are still in the nucleus, it is said to be *in situ*, or in the original location. Flow cytometry is a technique to determine the total amount of DNA within a cell. Although this is not a direct way to visualize chromosomes, it allows researchers, along with chromosome number, to determine genome size, that is, how much genetic material is present in a cell, which has implications during hybridization between species. Genome size varies enormously among angiosperms (see Chapter 7).

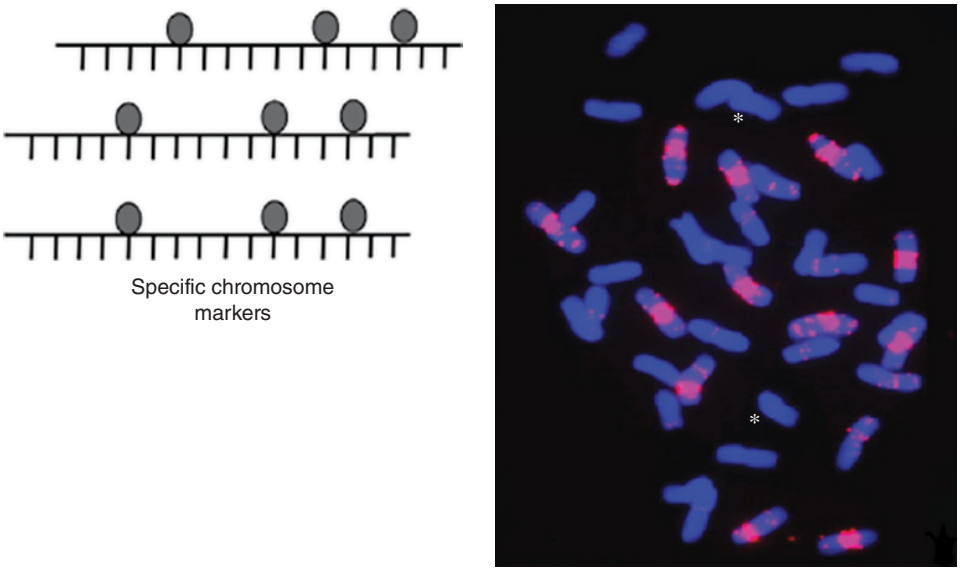


Figure 2.9. To the left is a diagrammatic representation of fluorescence in situ hybridization (FISH) for three chromosomes. The brightly-labeled loci on real chromosomes (right), show the physical location of a specifically-labeled DNA. The figure is modified from Zhang et al. (2015). (*See insert for color representation of the figure*.)

2.3.5. Mendelian Genetics and Biotechnology Summary

When a new gene is added to the genome by a biotechnological technique, it will then follow all the genetic processes described by Mendel. Most transgenes, genes that are placed into a genome by a biotechnological process, tend to act like a dominant gene in the genome. This new gene makes a functional protein, and therefore the functional protein is observed in the phenotype of the organism The genetically modified organism will often initially act like a heterozygote, due to the fact that one homologous chromosome has the locus, while the other does not (therefore, two different DNA sequences at that chromosomal location). The next generation involving this plant line would best be predicted using a monohybrid cross. One confounding possibility, it is also possible that multiple copies of the target DNA sequence will be added to genome in multiple independent locations, and therefore the phenotypic appearance of the trait may be controlled by multiple loci. In this case, let's say with two inserts, the transgenic phenotype would be best described using a dihybrid cross. The newly genetically modified organism must then proceed through sexual reproduction to produce seeds that represent the next generation. The new transgene will follow all the other genes through meiosis, and be part of a genetically unique gamete. The process of recombination can also involve transgenic DNA sequences, which can be a confounding factor as we try to predict the pattern of transmission of our new gene. Each different plant species will have a unique reproductive strategy, and therefore we must understand the multiple plant reproductive mating systems to predict how the new DNA will be passed to the next generation.

2.4. PLANT REPRODUCTIVE BIOLOGY

2.4.1. History of Research in Plant Reproduction

When it comes to sex, angiosperms have evolved many ways of doing it and indeed of doing without it. Sexuality in plants was first demonstrated experimentally over 300 years ago by a German botanist and physician Rudolph Jakob Camerarius. In his 1694 book *Epistolae de Sexu Plantarum* (*Letter on the Sexuality of Plants*), he identified the stamen and pistil as the male and female organs, and the pollen as the fertilizing agent. By the mid-1700s, the role of insects in pollination was well accepted, and in 1793 another German, Sprengel, provided elaborate details on the floral adaptations of 500 or more species to insect pollinators. Charles Darwin was also interested in pollination and plant mating systems from an evolutionary perspective, and one of his books, *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*, which was published in 1876, introduced the idea of self-incompatibility systems in plants. Plant mating systems have continued to fascinate botanists and geneticists since that time. Plant reproduction is clearly important to biotechnological improvements to agriculture, as it affects directly or indirectly the quality and quantity of all crop products.

The reproductive strategy used by each plant species will have an impact on the predictability of a transgenic phenotype added by biotechnology, and therefore it is important to have a good understanding of plant mating systems as transgenic organisms proceed through subsequent generations. The mating system will often influence whether a transgene will be maintained in the homozygous or heterozygous genetic state, which is important as a researcher makes predictions of the phenotype in the next generation.

2.4.2. Mating Systems

2.4.2.1. Sexual Reproduction. Traditional sexual reproduction is the best place to begin the discussion of plant mating systems. Seed production by sexual reproduction involves the transfer of pollen from an anther to the stigma of the pistil, followed by germination and growth of the pollen tube. The movement of nuclei in the pollen tube through the style to the embryo sac and

the union of functional male and female gametes complete sexual reproduction in plants. Pollination vectors, such as insects or wind, are responsible for the transfer of pollen, but mating systems determine if the pollen grain can germinate on a receptive stigma and penetrate the style. Mating systems are classified according to the source of pollen that is responsible for fertilization. Self-fertilization or selfing (also known as autogamy) occurs when the pollen that effects fertilization is produced on the same plant as the female gamete with which it unites. Cross-pollination or outcrossing (xenogamy) occurs when the pollen of one plant is responsible for fertilization of the female gamete of another plant.

The mating system of a plant species is also classified according to the relative frequency of self- versus cross-pollination in their seed production. There is a continuum of variation among species, ranging from complete selfing to obligate outcrossers, with those species demonstrating both often referred to as having a mixed mating system. Most crops have been bred and selected for selfing, but can also be outcrossed. This situation enables "true" seed to be produced by selfing in which the progeny is very genetically similar to the parent. "Homozygosity begets homozygosity." This situation also allows plant breeders to "shuffle" genomes from outcrossing when needed. The predominant mechanism of pollination for a species is an important factor in determining the breeding method used to develop the cultivar (see Chapter 3). For example, hybrid seed production is more readily accomplished in an outcrossing species than in a selfing species. The formation of homozygous lines occurs naturally in a self-pollinating species, but artificial self-sib-pollination must be practiced in outcrossing species to obtain homozygous genotypes. Both flower morphology and development, as discussed in more detail in the text, can influence rates of self- and cross-pollination.

2.4.2.1.1. Selfing (Autogamy) vs. Outcrossing (Xenogamy). Some plants have natural mechanisms that encourage self-pollination. One such mechanism in which pollination takes place while the flower is still closed is known as cleistogamy, and is a process that can occur even in self-incompatible species (Fig. 2.10). Homogamy, the synchronous maturation of stamens and stigma, also facilitates self-pollination.

The effects of repeated self-fertilization, first documented in maize by East and Shull at the turn of the twentieth century, have been confirmed for many crop species. Repeated self-fertilization will yield complete homozygosity in a few generations unless the heterozygous state is favored by selection. In a heterozygous diploid, the dominant allele can shelter recessive alleles that would be deleterious in the homozygous state. Self-fertilization quickly results in the segregation of lethal or sublethal types as homozygous recessives are produced. Further selfings rapidly separate the material into uniform lines, often called pure lines. Some of the the surviving lines may be characterized by reduced vigour and fertility, a condition known as inbreeding depression. If pure lines originating from different parental stocks are crossed together, hybrid vigor (i.e., heterosis) may be demonstrated. Outcrossing, thus, avoids the deleterious effects of inbreeding depression,

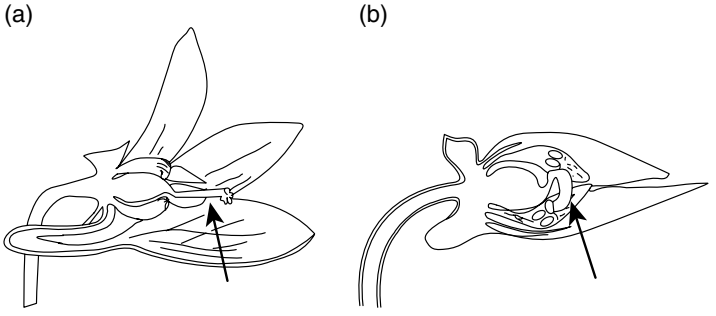


Figure 2.10. Cleistogamous flowers (b) are fertilized prior to the opening of sepals and petals, which ensures that the plant is self-pollinated. A non-cleistogamous flower is shown in (a). (Adapted from Briggs and Walters 1997.)

and promotes heterozygosity, genetic variability, and genetic exchange. Plants species have therefore evolved a wide variety of natural mechanisms that favor cross-pollination; scientists have needed to invent an alphabet soup to describe the myriad of mating syndromes observed. Several of these, including protandry, protogyny, chasmogamy, heterostyly, and imperfect flowers on monoecious, dioecious or polygamous plants, and incompatibility, will be discussed in somewhat greater detail later.

2.4.2.1.2. Sex Distribution within a Flower and within a Plant. Plants are the ultimate hermaphrodites—most species are bisexual with male and female organs together in one flower (also referred to as a perfect flower), but there are many ways in which sex organs are distributed within a flower, within a plant, and within a plant population. Some plants have separate male (staminate) flowers and female (pistillate) flowers on a single plant and are termed "monoecious" (e.g., maize). In other species, the male and female flowers occur on separate plants (known as dioecy), or can have a mixture of male, female, and perfect flowers on the same plants (termed "mixed polygamous"). Sex determination in such plants is under genetic control, with monoecy in maize, for example, under the control of a set of genes known as the tasselseed loci. A number of different mechanisms have been identified that establish the sexuality of dioecious plants, including the presence of heteromorphic sex chromosomes with males having XY and females XX chromosomes, or varying X –to –autosome ratios similar to that found in Drosophila. Even when both male and female organs occur in the same flower, the timing of sexual expression can vary. Sometimes, pollen is shed before the stigma is receptive in a process known as protandry, or a stigma can mature and cease to be receptive before pollen is shed (protogyny).

2.4.2.1.3. Self-Incompatability Genetic Systems. Many plant species have a genetic self-incompatibility (SI) mechanism that promotes outcrossing and is defined as "the inability of a fertile hermaphrodite seed plant to produce zygotes after self-pollination." SI mechanisms are estimated to occur in more than half of all angiosperm species. The effectiveness of SI in promoting outbreeding is believed to be one of the most important factors that ensured the evolutionary success of flowering plants, an idea first promoted by Darwin. It is a genetically controlled phenomenon, and in many cases the control is by a single locus known as the S locus. This locus often has up to several hundred alleles in some species. The SI mechanism promotes outcrossing by arresting "self" pollen tubes as determined by the genotype at the S locus (Fig. 2.11). SI is based on the ability of the pistil to recognize the presence of self-pollen and to inhibit the germination or subsequent development of self-related, but not genetically unrelated pollen. There are two types of SI mechanisms: gametophytic and sporophytic (Fig. 2.11); these differ in whether the haploid pollen genotype or the diploid pollen parent genotype determines the success of pollination. These are important traits for controlling pollinations and are much sought after in breeding programs.

2.4.2.1.4. Male Sterility. The ability to produce hybrid seed has been of fundamental importance to modern agricultural practice. "Hybrid vigor" has dramatically increased the yield in maize during the past 40 years. The genetic approach to the production of F_1 hybrid seed was made possible by the exploitation of various male sterility mechanisms. Male sterility refers to the failure of a plant to produce functional pollen by either genetic or cytoplasmic mechanisms. Cytoplasmic male sterility (CMS) is a maternally inherited trait that suppresses the production of viable pollen grains. It is a common trait reported in hundreds of species of higher plants. The CMS phenotype (female parent) is used commercially in the production of F_1 hybrid seed by preventing self-fertilization of the seed parent, in such crops as maize, sorghum, rice, sugar beet, and sunflower. The use of CMS lines as female parents also requires the introduction of nuclear fertility restorer genes from the pollen parent, so that male-fertile F_1 hybrids can be produced. Novel sources of CMS and fertility restorer genes are very important to plant breeders and the traits can be introduced via biotechnological means.

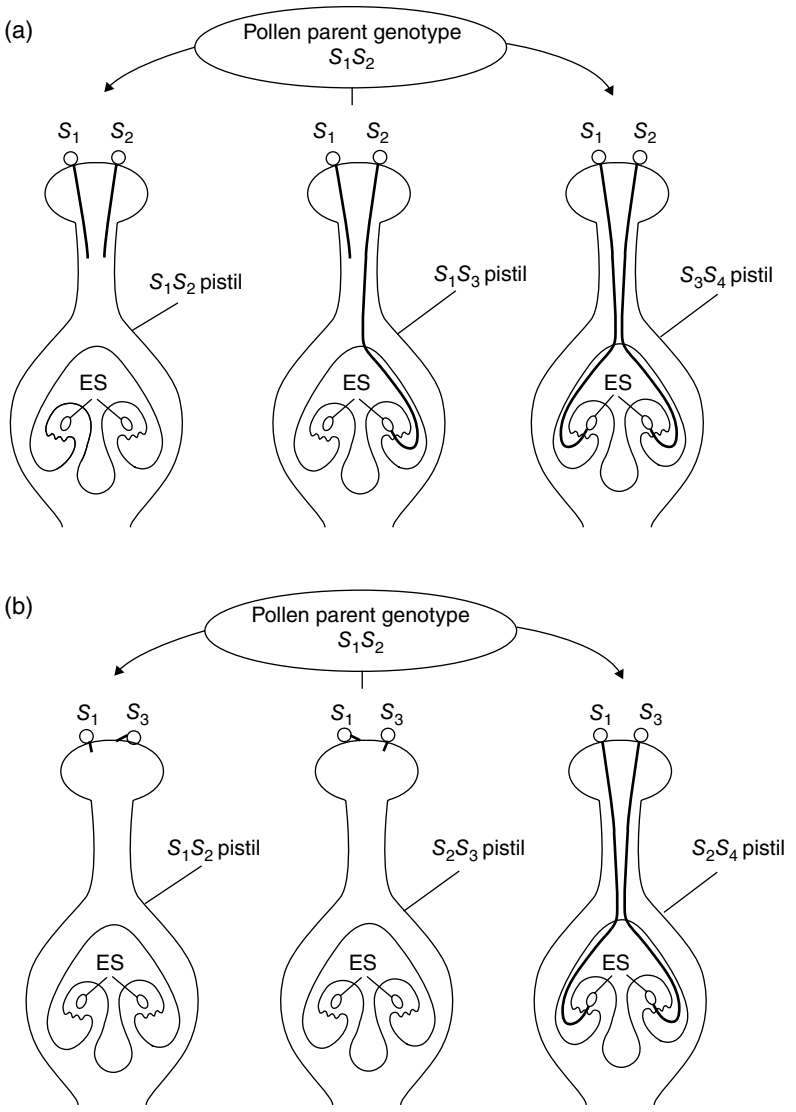


Figure 2.11. Self-incompatibility systems in plants may be gametophytic (a) or sporophytic (b). In gametophytic self-incompatibility, the pollen grain will not grow and fertilize ovules if the female plant has the same self-incompatibility (S) alleles. In sporophytic self-incompatibility, the diploid parent prevents germination of pollen grains that share an allele with the parent. (Adapted from Briggs and Walters 1997.)

2.4.2.2. Asexual Reproduction. Plants can also reproduce by asexual means, resulting in the multiplication of genetically identical individuals. An individual reproducing asexually is referred to as a clone and the process as cloning. Potatoes and cranberries are two plants that are primarily propagated by asexual reproduction. Asexual reproduction in seed plants can be divided into two main classes: vegetative propagation that can occur through plants parts other than seed (e.g., bulbs, corms, rhizomes, stolon, and tubers) and apomixis that can be defined as the production of fertile seeds in the absence of sexual fusion of gametes or "seeds without sex." Sexual fusion presupposes a reductional meiosis if the ploidy level is to remain stable. During apomixes, the embryo may develop from either an N (haploid) egg cell or from a 2N (diploid) egg cell. In the latter type known as agamospermy, a full reductional meiosis is usually absent and chromosomes do not segregate.

Another rare form of apomixis is that in which the embryo plant arises from tissue surrounding the embryo sac. These "adventitious" embryos occur, for example, in citrus crops.

2.4.2.3. Mating Systems Summary. Having discussed the three main modes of reproduction, selfing, outcrossing, and apomixis, we may now examine the advantages and disadvantages of different mating systems. One possible advantage of repeated self-fertilization is that well-adapted genotypes can be replicated with little change. A further advantage, especially in extreme or marginal habitats, where crossing between plants might be hazardous or fail all together, is that self-fertilization is an assured method of producing progeny. Outcrossing, on the other hand, avoids the deleterious effects of inbreeding depression, the main disadvantage of repeated selfing, and promotes heterozygosity, genetic variability, and genetic exchange. There are, however, costs to the plant, compared with selfers, as more biomass has to be employed in producing flowers, nectar, and so on. Other disadvantages to an obligate outcrosser are that if only one genotype is present in an area, the plant may not be able to reproduce sexually, or reproduction may be rendered uncertain or unlikely by environmental factors. With outcrosssing, each generation produces new variability, and although most progeny may be fit and well-adapted, some progeny may be less fit and constitute "genetic load" to the population. The third method of reproduction—apomixis—facilitates the production of a large number of well-adapted plants of the maternal genotype with little or no genetic load. Apomixis offers the possibility of reproduction by seed in plants with "odd" or unbalanced chromosome numbers, such plants being unable to produce viable gametes at meiosis and likely to be totally or partially seed-sterile. Seed apomixis, for example, provides all the advantages of the seed habit (dispersal of propagules and a potential means of survival through unfavourable seasons). Apomicts are often of polyploid and hybrid origin and therefore this reproductive mode can potentially serve as a means of preserving high heterozygosity. Apomixis, like selfing, would also appear to be important at the edge of the range of a species allowing populations to persist in areas in which various factors may limit or exclude the possibility of sexual reproduction. Given that all three reproductive modes have advantages and disadvantages depending on environmental circumstances, it is not too surprising to learn that plants often have highly flexibile mating systems, reproducing by several means, rather than just relying on a single reproductive mode.

The mating system of a plant species will influence the way in which the genetic diversity present in the species is distributed within and among its populations—that, its population structure. In outcrossing species, higher levels of genetic diversity are found within populations than among populations. The opposite is true for predominantly selfing species where greater differentiation among populations is expected. Knowledge of a plant's mating system is important in conservation of its genetic diversity in a seed genebank, or for efficient screening of populations of wild species as source of traits for crop improvement in plant breeding programs. More populations of a selfing species would be needed in order to capture the true diversity of a species.

2.4.3. Hybridization and Polyploidy

Although we think of species as discrete and static breeding entities, examples can be found throughout the angiosperms where different species have the capacity to cross with another. Plants are champions at interspecific hybridization. Hybridization, or the process of sexual reproduction between members of different species or biotypes within a species, produces plants that have genetic material from both parents. In most cases, the initial hybridization event results in hybrid plants that are haploid for each genome or, in other words, have a single homologous chromosome from each parental chromosome set. As homologous chromosomes are normally paired during metaphase I, the presence of only one of each homologous chromosome pair can disrupt normal meiotic function. In fact, most of the gametes produced in hybrids are abnormal, leading to sterility to reduced viability of pollen or eggs in the hybrid plant. Although hybrids can be made from the crossing of many different species, hybridization of normal haploid gametes rarely generates plants that are fully fertile.

In some cases, sex cells are produced that have more than just one of each homologous chromosome. Nondisjunction, when homologous chromosomes fail to separate during meiosis, sometimes generates gametes that have complete sets of chromosomes from the parent species, called unreduced gametes. If two unreduced gametes fertilize one another, the resultant hybrid would have the complete genome of each parental species. In this case, meiosis can function normally, and the hybrid plant may represent a new species with a unique chromosome number. Species that contain multiple genomes or multiple sets of chromosomes beyond the diploid level are called polyploids. Again, among the myriad of organismal types, plants are champions at polyploid production—and indeed many plant species are polyploids.

Polyploidy may arise in two ways: by the doubling of a homologous set of chromosomes (autopolyploidy) or by combining two complete sets of chromosomes from genetically different parent plants (allopolyploidy). An autotetraploid contains four sets of homologous chromosomes, and pairing between the four homologous chromosomes is often irregular with chromatids showing random segregation during gamete formation. In an allotetraploid on the other hand, the parental chromosomes in each of the two sets of homologous chromosomes tend to pair with each other as they would in the parental plants, thus contributing to the stability and fertility of such plants. Several natural allopolyploids are known and several have been created in the plant breeding field.

Hybridization is an important process that has occurred in the development of many of our agricultural crops. Many polyploidy crop plants have been produced by either the combination of unreduced gametes or the doubling of the chromosomes after hybridization of haploid gametes. Canola (*Brassica napus*), which is used for vegetable cooking oil, is composed of the complete genomes of two different species (*Brassica rapa*, genome AA and *Brassica oleracea*, genome CC); and similar polyploid origins have been confirmed for two other mustard crops *Brassica juncea* and *Brassica carinata* (Fig. 2.12). Bread wheat, *Triticum aestivum*, was produced from the hybridization between three different species. In this case, each progenitor species donated their complete diploid genome (AA, BB, DD genomes, respectively) to making a species with three complete sets of chromosomes and a very large "new" wheat genome (AABBDD).

Polyploidization is undoubtedly a frequent mode of diversification and speciation in plants. Most plants have undergone one or more episodes of polyploidization (i.e., increase in the whole DNA complement beyond the diploid level) during their evolution (Soltis et al. 2004). Hybrid speciation is another important phenomenon. Interspecific hybridization and subsequent introgression of the portion of the genome of one species into that of another (Fig. 2.13) have often been recognized as a source of genetic variation and genetic novelties, and in some cases successful hybridization events have promoted rapid speciation radiation. The complexities of plant genetics

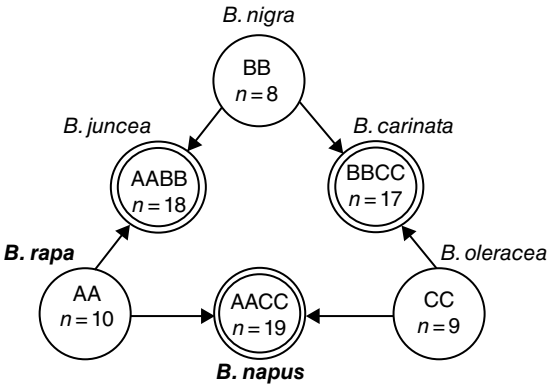


Figure 2.12. Triangle of U (1935) shows the relationships between several diploid and polyploidy crop species within the *Brassica* genus.

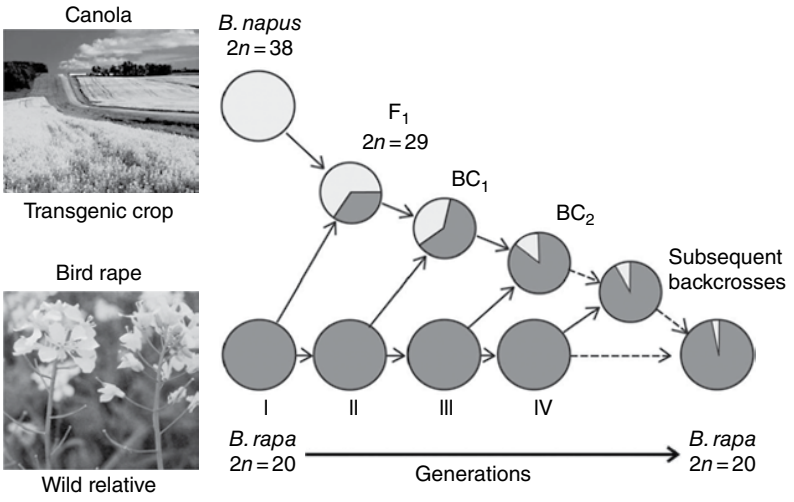


Figure 2.13. Hybridization and genetic introgression between closely related species allow for the incorporation of genetic material from one species to another.

can be traced to reproductive biology and mating systems in plants, which is an area of research that is very active and dynamic.

2.4.4. Mating Systems and Biotechnology Summary

Depending on the mating system of the target plant species, a new transgenic phenotype will be either straightforward or very complicated to maintain in a predictable fashion in subsequent generations. For self-compatible species, the primary transgenic plant will simply self, and homozygous individuals will be selected and the phenotype will be fixed for future generations. For self-incompatible species, the primary transgenic plant must be crossed with a non-transgenic individual with different self-incompatibility alleles. In this case, it will take a concerted breeding strategy to maintain the transgene, and the trait will often be maintained in the heterozygous state. It may be challenging to produce homozygous individuals with self-incompatibility, as each generation requires outcrossing with a different genotype. Polyploid plant species will present unique challenges, especially if the transgenic line will be crossed with progenitor species to acquire other traits through hybridization. The genome in which the transgene was incorporated will determine the success of such a breeding strategy. The mating system of the plant will have an impact on the predictability of the transgenic phenotype, and therefore must be considered when selecting a particular plant type for genetic transformation.

2.5. CONCLUSION

After biotechnologists introduce or manipulate genes in plants, if all goes as expected, the new genes should be part of the genomic fabric and behave like “normal” plant genes. Therefore, they should follow the laws of Mendelian genetics and be passed on to future generations like other genes of the particular species. Therefore, it is important for the plant biotechnologist to understand basic Mendelian genetics and plant reproductive systems to properly plan as the recombinant plant line is taken to future generations. Transgenes also become part of breeding programs, which is why understanding the fate of transgenes in new plant cultivars is important—the subject of Chapter 3.

LIFE BOX 2.1. RICHARD A. DIXON

Richard A. Dixon, Distinguished Research Professor, Department of Biological Sciences, University of North Texas; Member, US National Academy of Sciences; Fellow, National Academy of Inventors.



Rick Dixon relaxing at a faculty retreat, Quartz Mountain, Oklahoma (May 2007). Courtesy of Rick Dixon.

I first became interested in plant natural products as an undergraduate at Oxford. I was reading Biochemistry, and the course was quite heavily weighted towards physical biochemistry, an area I found hard because of my lack of mathematical prowess. Faced with the choice of either whole animal physiology or plant biochemistry as an elective, I jumped at the latter, a decision that determined the future course of my career. I had been excited by organic chemistry at an early age, and was fascinated to learn how plants "do" organic chemistry during the synthesis of natural products and lignin. This was before the era of molecular biology, and our understanding depended mostly on the results of *in vivo* labeling studies coupled with *in vitro* enzymology. I always remember my first lecture from Vernon Butt, in which he outlined current views on how the monolignol units of lignin are formed. It all seemed so beautiful and logical, although my group and others were later to show that it is actually more complex than envisaged at the time. This new understanding had to wait until we had the necessary genetic and genomic tools.

I decided to stay on in the Botany School at Oxford to work on my D.Phil. with Keith Fuller. Keith had suggested a project on galactomannan mobilization in alfalfa; but when I returned from the summer vacation to start this project, we discovered that four papers, reporting essentially everything we were planning to do, had just appeared in the literature. Keith suggested I might instead look at how plants make bioactives in cell culture. I was disappointed at being "scooped" on my planned project (although better early than later!), and did not realize at the time that agreeing to the back-up plan was the defining moment in my career. Using the isoflavonoid phytoalexin phaseollin from bean as a model, I established conditions for turning on isoflavonoid metabolism in cell cultures. When Chris Lamb joined the lab as a postdoc we set up a collaboration that lasted nearly 20 years, in which we used the phytoalexin induction system as a model for studying microbially induced gene expression in plants using the new tools of molecular genetics.

After 2 years of postdoctoral work in Cambridge and 9 years of teaching and research at the University of London, I moved to become director of the newly formed Plant Biology Division at the Noble Foundation in Ardmore, Oklahoma, in 1988. During the first 8 years of my tenure at Noble, I continued to work primarily on plant–microbe interactions. The Noble Foundation's major mission is to assist farmers and ranchers reach their production goals through basic and applied science and demonstration; and during the previous years, I had hired a number of excellent principal investigators in the plant–microbe interaction field. I therefore decided to move away from the plant–microbe focus and concentrate my research on those natural product pathways that impacted forage quality, the health of ruminant animals, and human health. This

was another decision, dictated by circumstances, that has paid dividends. The work I initiated on the biosynthesis and metabolic engineering of lignin and proanthocyanidins has been rewarding as basic science, has moved to successful commercialization or reduced lignin alfalfa through a long-term research collaboration with Forage Genetics International, and has had important implications for plant metabolic engineering in relation to lignocellulosic bioenergy crops (lignin) and human health (proanthocyanidins). This is certainly more than I envisaged when I first decided that the plant- microbe field was too crowded and that quieter pastures might profitably be grazed! I am still working on these projects since my move to the University of North Texas in 2013, where I have also initiated new collaborative work on bio-based products from lignin.

Based on my personal experiences, my advice to young scientists would be to always stick with what you are passionate about, always try to work with people who are smarter than you are, and never turn down opportunities to adapt your program to emerging applications. It is also critical to get away from the lab and clean out your brain (regularly!). I have had a passion, since the age of 10, for studying, collecting and cultivating cacti. I also love hiking, particularly in mountains. The photograph shows me indulging both of these passions in the Quartz Mountains of Southwestern Oklahoma (although I have to admit that this was during a short break at a faculty retreat!).

LIFE BOX 2.2. MICHAEL L. ARNOLD

Michael L. Arnold, Distinguished Research Professor of Genetics, University of Georgia.



Mike Arnold with *Iris nelsonii*; Vermilion Parish, Louisiana. Courtesy of Mike Arnold.

From Whence I Come

In regard to my career as an evolutionary biologist, I start the clock with the Fall, 1975 semester of my freshman year at Texas Tech University. During this time period, I fell in love with research science—sometimes to the detriment of my participation in classes! My initial plan was to work with a parasitologist who specialized in organisms dug from the rotting remains of farm animals. However, this professor stood me up for several scheduled meetings, and so I turned instead to a plant evolutionary biologist, Professor Raymond Jackson, and an animal evolutionary biologist, Professor Robert Baker, as my first two mentors. Their patience and encouragement helped me to not only finish the lab work for several research projects but also to see the research published in scholarly journals. This taught me the love of discovery and creation—discovery of facts about the natural, evolving world and creation of word pictures in order to explain what had been discovered. The careful tutelage of my mentors gave me the understanding of how to pursue research projects. Because my earliest training was in both botany and zoology, it has been natural for me to emphasize tests for common evolutionary patterns between plants and animals that may reveal common underlying processes. This emphasis is reflected both by the breadth of organisms on which my students, post-doctoral associates and I have worked (everything from fruit flies to fungi and fruit bats to Louisiana Irises) and the synthetic treatments we have

produced—for example, the three books *Natural Hybridization and Evolution*, 1997; *Evolution Through Genetic Exchange*, 2006; and *Reticulate Evolution and Humans—Origins and Ecology*, 2009.

(Re)Turning to Plants

Though, as indicated earlier, my colleagues and I have examined many types of organisms, 27 years ago I did make a decision to focus most of my research efforts on plant taxa. Several factors led to this decision, two of which related to my earliest training in evolutionary botany and zoology. I had learned quickly, that testing many of the hypotheses in which I was interested—especially those associated with the processes of genetic exchange, speciation and adaptation—required taxa that would allow a dual approach of experimental manipulations and surveys of natural populations. Most plant and animal groups (and for that matter, many bacterial and viral assemblages) provide opportunities to examine naturally occurring populations for the purpose of estimating evolutionary processes such as genetic exchange via introgressive hybridization and/or horizontal gene transfer. However, few animal clades allow the type of direct assessments possible in studies of plant species (e.g., through reciprocal transplantations into both experimental and natural environments). In addition, my interest in testing the descriptiveness of the web-of-life metaphor (i.e., that emphasizes the importance of genetic exchange in the evolution of organisms) led me to choose plants over animals. Thus, evolutionary biologists consider plants to be paradigms of such processes as introgressive hybridization, hybrid speciation and adaptive trait transfers.

Has Our Work Affected Plant Biotechnology?

I believe that the work carried out by my colleagues and myself has impacted the field of plant biotechnology in several ways. However, all of the effects from this work can likely be traced back to our emphasis on studies of population level phenomena. In the early 1990s, when we began our research into reticulate evolution, plant evolutionary biology was characterized by systematic treatments (i.e., studies that defined the relationships of species). Many decades had passed since the appearance of the wealth of publications by such workers as Edgar Anderson and Ledyard Stebbins on the population-level phenomena associated with genetic exchange between plant lineages. With few exceptions—for example, see many publications of Verne Grant and Don Levin—studies of plant evolution since Stebbins' and Anderson's time had emphasized pattern over process. In contrast, our work was designed to emphasize process over pattern. For example, we have asked how the processes of introgressive hybridization, hybrid speciation, lateral exchange, and adaptive trait transfer have affected the evolutionary patterns reflected in present-day biodiversity. This process-over-pattern focus has led to the application of our findings by plant biotechnologists, particularly when they are considering the effect that gene exchange might have on development and control of genetically engineered plants. One example of this can be seen in the interest that we have generated by highlighting the observation common to the vast majority of hybridizing plant and animal taxa (as well as for those organisms exchanging genes via viral recombination and lateral exchange), that hybrid genotypes demonstrate a range of fitness estimates that are often affected by the environment. This key observation leads to an array of expectations concerning the challenges faced in forming hybrid lineages—under both natural and experimental conditions. Furthermore, the observation of a wide range of hybrid fitness should also lead to caution during the generation of predictions concerning the effects on natural ecosystems from the introduction of genetically engineered plant lineages.

To Where Are We Going?

I am reminded of the Old Testament mandate that states that prophets, once proven inaccurate, were to be stoned. In that context, I offer the following suggestion concerning one direction I believe studies of genetic exchange (of which I do consider myself a student) and plant biotechnology (of which I do not) should be progressing. The analyses of genetic exchange, across all taxonomic categories,

are entering an exciting phase. The definition of the genomic architecture of related organisms allows the dissection of the causal factors that affect the transfer of specific loci. Given such information, it is possible to state with some certainty, which loci are prevented and which loci are facilitated in their transfer between organisms belonging to divergent evolutionary lineages. However, a more difficult, and much more significant, inference is needed. Specifically, it is necessary to define the "why" behind a transfer (or lack of transfer). In other words, what is the specific effect on the organism that causes either an increase in the fitness of hybrid genotypes (leading to genetic transfer) or a decrease in the fitness of hybrid genotypes (resulting in no transfer) when certain combinations of loci are present? With the advent of next- and next-next-generation sequencing, we are indeed beginning to dissect the "whys" of transfer events. For example, as our own species spread out from an African center of origin, we encountered related species (e.g., *Homo neanderthalensis*) and received via introgressive hybridization alleles that now allow our species to inhabit a variety of habitats including extreme elevations and different levels of UV radiation. Likewise, in our own research, we have detected the influence of introgression of alleles giving rise to flood tolerance in a species previously adapted only to drier soils. The degree to which we are able to address and answer why transfers have occurred will determine how well we will be able to test hypotheses concerning such fundamentally important processes as (a) the effect of genetic exchange on hybrid lineage formation and the transfer of adaptations and (b) the impact of genetic exchange between genetically engineered plants and wild relatives on both crop production and natural ecosystems.

REFERENCES

Briggs D, Walters SM (1997): *Plant Variation and Evolution*, 3rd ed. Cambridge University Press, Cambridge, UK.

Camerarius RJ (1694): Epistola ad M.B. Valentini de sexu plantarum. In *Ostwald's Klassiker der exakten Naturwissenschaften*, No. 105, 1899. Verlag von Wilhelm Engelmann, Leipzig.

Darwin CR (1876): *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. John Murray, London.

Goss J (1824): On the variation in the colour of peas, occasioned by cross-impregnation. *Trans Hort Soc Lond* **5**:234.

Knight TA (1799): An account of some experiments on the fecundation of vegetables. *Phil Trans Roy Soc Lond* **89**:195–204.

Mendel G (1866): Experiments in plant hybridization (translation). In *Classical Papers in Genetics*, Peters JA, (ed). 1959. Prentice-Hall, Englewood Cliffs, NJ.

Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate J, Madrodiev E (2004): Recent and recurrent polyploidy in Tragopogon (Asteraceae): cytogenetic, genomic, and genetic comparisons. *Biol J Linn Soc* **82**:485–501.

Sprengel CK (1793): *Das entdeckte Geheimniss der Natur im Bau und in der Befruchtung der Blumen [The Secret of Nature in the Form and Fertilization of Flowers Discovered]*. Berlin: Bei Friedrich Vieweg dem ældtern.

U N (1935): Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn J Bot* **7**:389–452.

Zhang H, Li G, Li D, Gao D, Zhang J, Yang E, Yang Z (2015): Molecular and cytogenetic characterization of new wheat – *Dasypyrum breviaristatum* derivatives with post-harvest re-growth habit. *Genes* **6**:1242–1255.

CHAPTER 3

Plant Breeding

NICHOLAS A. TINKER AND ELROY R. COBER

Agriculture and Agri-Food Canada, Ottawa, Canada

3.0. CHAPTER SUMMARY AND OBJECTIVES

3.0.1. Summary

Breeding "re-shuffles and re-deals" thousands of genes at a time by hybridizing parental plants, then attempts to pick a new "winning hand" by selecting on traits (and genes) of interest. Successful plant breeding requires the informed choosing of parents, and the application of large-scale selection strategies that are designed based on principles of Mendelian genetics, probability, and statistics. In practice, when biotechnology is used, it is nearly always combined with conventional plant breeding for crop improvement by moving novel genes into appropriate genetic backgrounds.

3.0.2. Discussion Questions

1. Describe how plant breeding is both an art and a science.
2. Is seed color a qualitative or quantitative trait?
3. List six factors that can affect the distribution of quantitative trait phenotypes that will appear in a given population.
4. What proportion of plants in an F_6 generation will be heterozygous at a given locus?
5. What is the probability that five segregating loci will all be homozygous in the F_6 generation?
6. What is the difference between a landrace and a pure-line plant variety?
7. The pedigree method and the single-seed descent (SSD) method are two strategies for developing pure-line varieties. List some factors that might influence your choice of one versus the other.
8. Using the terms "homozygous or heterozygous," and "homogeneous or heterogeneous," describe each of the following: (a) a modern maize hybrid, (b) a synthetic alfalfa variety, (c) a mass-selected population of maize, (d) a landrace of wheat, and (e) a modern variety of wheat.

*This chapter is the work of the Department of Agriculture and Agri-Food (Canada) (AAAF).

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

3.1. INTRODUCTION

Plant breeders enjoy quoting a famous US president, who wrote that "The greatest service which can be rendered any country is, to add a useful plant to its culture."[1] Whether or not you agree, it must be acknowledged that the creation of new and better plant varieties is among the most useful and visible outcomes of biotechnology. Whether it is a noble service might depend on whether you do it for fun, for profit, or for the good of humanity, but most breeders will confess to all three motives. Plant breeding has been credited with helping to triple the productivity of modern agriculture, and it has been a fundamental part of international humanitarian achievements (Hoisington et al. 1999). But you do not have to "think big" to be excited about plant breeding. Admire the colors on your next plate at dinnertime, taste the subtle flavors in your next bite of fruit, feel the strength and softness of your cotton shirt, or smell your favorite rose—these characteristics are all derived from unique characters of different plant varieties. You spend less of your income to eat a much better variety of foods than your ancestors did, largely because of plant domestication and breeding. In future, you might live longer or healthier because of the varieties of plants used to make your breakfast cereal. What fun it would be to create those varieties, or even just to understand how they are created!

Plant breeding is a skill that requires advanced learning and practical experience. Many universities and corporations worry that trained plant breeders are becoming scarce in relation to ongoing demands. Most of the modern concepts in biotechnology that are introduced in this book can be viewed either as enhancements to plant breeding, or as innovations that can be rendered useful only through plant breeding. The most groundbreaking achievements in biotechnology still need to be packaged in plants that are productive, disease-free, tasty, and nutritious. These qualities depend on the coordinated expression and complex interactions of thousands of plant genes and gene products. We have learned many things about many genes, but we may never know enough to fine-tune all of the genes required to make a plant variety that is adapted and competitive under modern agricultural production practices. Thus, we continue to depend on plant breeding as the cornerstone of commercialization and technology transfer.

Plant breeding was described by Nikolai Vavilov,[2] the famous Russian scientist, as "evolution directed by man." Thus, the job of a plant breeder is to replace natural selection with artificial selection, such that combinations of traits can be assembled into plant varieties that would not otherwise be found in nature. While correct, this definition hides many of the dimensions in which a breeder must work to produce successful plant varieties. There are two primary interventions that a breeder makes: the deliberate hybridization of specific parents and the selection (or elimination) of progeny. This seemingly simple iterative process is elaborated by many factors: knowledge of what traits are important, knowledge of genetic control, knowledge of how environment affects traits, and knowledge of strategies to reduce the sheer number of progeny that must be examined. On top of this, a breeder must be a communicator, a team builder, an extension worker, an expert in commercialization, and a specialist in legal, ethical, and social issues. Plant breeding is often

[1] From Memoir, Correspondence, and Miscellanies, from the papers of Thomas Jefferson, edited by Thomas Jefferson Randolf, 1829, p. 144.

[2] Vavilov and Lysenko: In the plant sciences, Russia has produced one of the most influential plant scientists as well as one of the most notorious. Nikolai Ivanovich Vavilov (1887–1943) is credited with several important discoveries in genetics, including the demonstration that the center of diversity of a plant species is an indication of its center of origin. He also assembled one of the largest and most diverse collections of plant germplasm in the world, now housed at the Vavilov Institute of Plant Industry in St. Petersburg. Trofim Denisovich Lysenko (1898–1976) was an experimentalist who claimed the discovery of many agricultural methods that now seem absurd. To his credit, he also studied some phenomena such as vernalization that we now recognize as important physiological mechanisms. Unfortunately, his claims of rapid and phenomenal success were popularized to the extent that he garnered great political influence in Stalin's Soviet Union. When he was put in charge of the Academy of Agricultural Sciences of the Soviet Union, he was able to silence or imprison his critics, including Vavilov, who died in prison in 1943.

described as being an art and a science. While there are deterministic principles to discover and apply, there is often more than one acceptable result, and more than one way to achieve the same result. Plant breeders sometimes claim to recognize another breeder's "handiwork" by the way a variety looks in the field, and often they find that the most efficient use of time and resources is to walk through a field and identify plants that "just look right."

While studying the topic of plant breeding, you might think of numerous analogies that help you conceptualize the process. For example, you might draw a parallel between a good plant variety and a favorite song; they are both dependent on many subtle characteristics; and although their quality may be widely acknowledged, appreciation of this quality is varied. In some ways, breeding is also similar to the iterative trial-and-error process that investors use to build strong and diverse investment portfolios, and principles of genetic selection have even been applied with great success to areas such as this. But the processes of genetic recombination and gene expression are unique to DNA-based organisms, and no analogy can completely replace the concepts that must be learned to become a successful plant breeder.

This chapter introduces some fundamental concepts of plant breeding, and describes some generic strategies that are typically used to breed plant species that have a variety of mating systems. In the "real world," every plant species presents unique challenges and opportunities, and it is beyond the scope of this chapter to discuss strategies used for specific crops. The emphasis in this chapter has been placed on describing the underlying concepts of plant breeding, which will help you understand and appreciate literature that is more detailed or specific. You are encouraged to look at some of the references listed at the end of this chapter to see how breeding is typically applied in plant species that interest you.

3.2. CENTRAL CONCEPTS IN PLANT BREEDING

Prior to reading this chapter, the reader should understand the principles in Chapter 2 covering basic Mendelian genetics and plant reproduction. The concepts introduced in this section will build on that knowledge. The following paragraphs introduce key concepts that collectively determine most of the decisions and strategies of a breeding program.

3.2.1. Simple vs. Complex Inheritance

Chapter 2 introduced Mendelian genetics—undeniably, the most important concept that a breeder must understand. The discovery of Mendelian principles was made in a plant species (pea) using traits that might be important in a pea breeding program (color, height, and starch content). These traits are considered *qualitative* (having discrete values such as green or yellow and tall or short) and *monogenic* (controlled by single genes). Such traits are also described as showing *simple inheritance*. However, many other traits that a plant breeder would consider—such as fruit weight, maturity date, and grain yield—are *quantitative* (measured on a continuous scale) and *polygenic* (controlled by many genes). Such traits are also described as showing *complex inheritance*.

Figure 3.1 illustrates how a quantitative, polygenic trait can still have underlying Mendelian inheritance. In this illustration, the size of a melon fruit is determined by the type of alleles that are present at two different genetic loci. This type of mathematical simplification is commonly used to develop or test models that can help explain the numbers of genes and the types of gene action that are involved in the expression of quantitative traits. Although these are mathematical assumptions, models such as this one can often approximate underlying biological phenomena. For example, the "capital" alleles in Figure 3.1 could represent gene promoters that trigger higher expression of fruit development factors, and the "small" alleles are less effective versions of these gene promoters.

The distinction between simple and complex inheritance is a common source of confusion. We say that green versus yellow is a simple monogenic trait, because it is often determined by one of

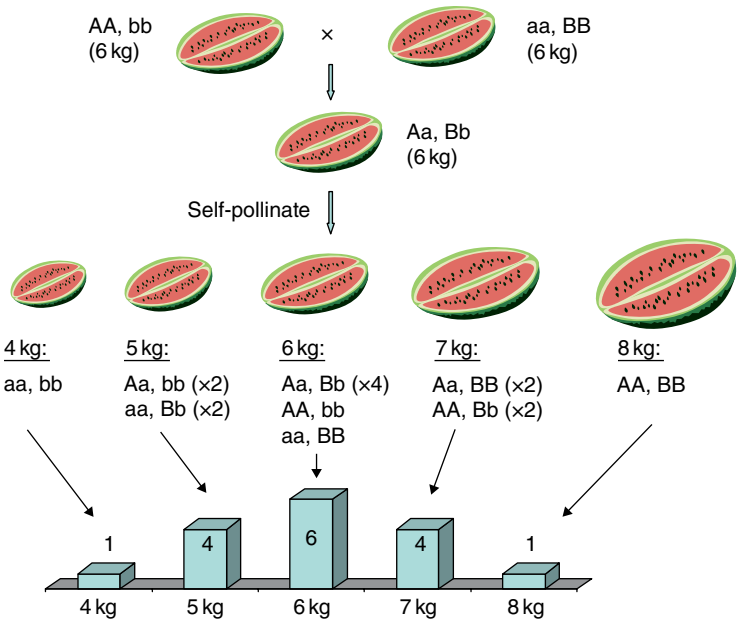


Figure 3.1. A hypothetical melon-breeding scenario that illustrates quantitative inheritance. Alleles at two loci (A and B) are represented by lower versus uppercase letters. Assume that the allele represented by the capital letter increases melon weight relative to the small allele, such that the average weight of melons (in kg) produced by a variety is determined by two times the total number of capital alleles plus the number of small alleles. Nine different genotypes will be present in the F_2 generation. If the two loci are not linked on the same chromosome, their expected segregation ratios will be as depicted in Figure 2.5. However, because some different genotypes result in the same melon size, only five sizes of melons will be produced, as shown. The expected proportions of each melon size in a random F_2 population are depicted by the histogram at the bottom of the figure.

two alternate alleles at a single genetic locus. But there are probably numerous other genes that might influence the intensity of the green or yellow color, and there are probably other gene loci that could mutate to block the production of the green color. These other differences may or may not be noticeable, and they may or may not be present within a specific germplasm. Thus, in some populations, seed color could be a polygenic trait. Indeed, a pea breeder might make a cross between a yellow-seeded variety and a green-seeded variety, and then try to select progeny that had even greener seeds than those of the original green parent. Why? Because the yellow-seeded variety might contain alleles at loci other than the primary seed color locus that are capable of enhancing the green color in the presence of the green alleles at the primary locus. Thus, seed color can be either qualitative and monogenic (simple inheritance), or quantitative and polygenic (complex inheritance) or both, depending on circumstance and on the germplasm being investigated.

3.2.2. Phenotype vs. Genotype

An important term used throughout this book is "phenotype," which simply means "what something looks like." We often speak about the phenotype of a specific trait, in which case it takes on units of measurement. For example, the phenotype of a quantitative trait such as seed weight in wheat might range between 30 and 80 mg. The term phenotype is also used to distinguish what a plant looks like from its genotype (what genes are present). Interactions between the genotype of a plant and the environment where it grows can lead to different phenotypes, a phenomenon called genotype-by-environment interaction (or G×E for short). For example, two wheat varieties could have a similar phenotype in a

disease-free environment but show distinct disease-resistant or susceptible phenotypes when grown where fungus spores are present. If we could identify or control all the unpredictable effects of error and environment, then the phenotype of a plant (P) would be equal to its genotype (G) plus the G×E interaction, plus the effects caused by the environment (E). Most of the fancy equations that you will see in plant breeding books (e.g., Allard 1999) are derivations of the following basic formula:

$$\mathrm{P} = \mathrm{G} + \left(\mathrm{G} \times \mathrm{E}\right) + \mathrm{E}$$

Here, it must be remembered that G and G×E are themselves made up of all the individual effects of genes present in the plant. Genes can also interact with each other, so more complex versions of this equation can be written to represent this. However, the above equation captures the essence of plant breeding, and reminds us that a plant that does well in one environment may not do well in a different environment.

The equation above refers to the genetic or phenotype values of a single plant or plant variety. However, breeders work with populations of many plants, and spend much of their time sorting through populations that contain variable genotypes. In fact, genetic variability is the key to creating new plant varieties through artificial selection. To summarize the amount of variation in a population, we calculate a statistic called the variance. The basic breeding equation can also be written to describe a population of plants in terms of phenotypic variance (V_P), genetic variance (V_G), and environmental variance (V_E), such that

$$V_P = V_G + V_E$$

Here, we have omitted variance attributable to G×E for simplicity. It is imperative for any breeder to understand the relative proportion of genetic variance that contributes to phenotypic variance for a given trait. This concept is formalized using the term "heritability" (H), which, in its simplest form, is measured as follows:

$$H = \frac{V_G}{V_P}$$

Since V_P is always greater than or equal to V_G, the heritability of a trait can range from 0 to 1. If H is equal to one, then all variance is caused by genetic effects, and the breeder will be very successful at selecting better plants. Such is the case for the imaginary melon trait illustrated in Figure 3.1. However, if H is zero, then V_G must also be zero, and there is no possibility of selecting plants that are genetically superior because all variation is environmental. Most traits that breeders work with show intermediate levels of heritability, between 0 and 1. A trait with low heritability requires more effort to measure since multiple environments are needed to accurately measure the trait compared to a trait with high heritability which might be accurately measured in only one environment.

3.2.3. Mating Systems, Varieties, Landraces, and Pure Lines

The fundamental output of agricultural plant breeding is to produce a plant variety, which is sometimes referred to as a cultivar (i.e., cultivated variety). However, the genetic makeup of a variety, and the way in which it is produced, maintained, and released depends critically on the type of mating system found in the species to which the variety belongs. Many plant species can tolerate self-pollination (or self-fertilization), and some of the most important crop species (including most grain and oilseed crops) are naturally self-pollinated. An important exception is maize (corn), which can tolerate self-pollination, but is normally cross-pollinated (or cross-fertilized). Other plants cannot tolerate self-pollination, and have specific genetic mechanisms to prevent this (see Chapter 2). Plants

that normally cross-pollinate are subject to continual recombination and selection after varieties are released, and thus strategies for breeding and variety release can be quite different from those used in self-pollinating species.

For plant species that normally cross-pollinate, we often assume that mating occurs at random. In reality, this is seldom the case because plants that are near to each other are more likely to pollinate each other. Nevertheless, the assumption of random mating allows the development of theories that often give good approximations of reality. The most important theory regarding random mating is the Hardy–Weinberg law, which predicts the frequency of genotypes that will occur according to the frequency of alleles. Assume that there are two alleles, "A" and "a," at a given locus, and that the alleles are at frequencies p and q, where p must equal $(1-q)$. The law states that the frequencies of genotypes, as represented below, can be predicted as follows:

$$\mathrm{AA}:\mathrm{Aa}:\mathrm{aa} = p2:2\ \ pq:q2$$

An important property of this law is that these frequencies are achieved after just one generation of random mating (the proof of this theory is shown in many textbooks). An important application of this theory is to identify whether random mating is occurring, or if other factors such as selection or mixing of populations (immigration/emigration) are occurring.

Plant species that are highly self-pollinated usually exist in a homozygous state (i.e., alleles exist in identical pairs at most loci). To understand why, consider what happens when a hybrid is formed, through either a chance pollination or a deliberate hybridization by a breeder. Figure 3.2 shows a cross between two homozygous genotypes. The product of this mating (a hybrid) will be heterozygous at any locus that differs between the parents, and all progeny will be identical. However, a mixture of genotypes will exist in the F_2 generation and beyond. Each generation of selfing reduces the level of heterozygosity by 50%, such that the proportion of homozygotes (P_{homo}) at a particular locus in generation F_X can be predicted as follows:

$$P_{\mathrm{homo}} = 1 - \frac{1}{2}(X-1)$$

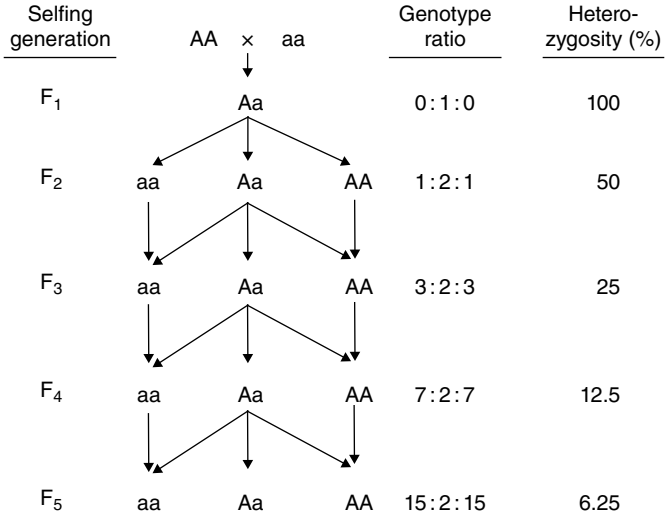


Figure 3.2. In repeated self-pollination with no selection, the level of heterozygosity is reduced by one-half with each selfing generation. This is because only half of the progeny from the heterozygous genotypes will still be heterozygous, while all the progeny from homozygous genotypes will be homozygous. Thus, the population gradually approaches complete homozygosity, but with a mixture of homozygous genotypes.

This prediction applies to a single locus, and to the average level of homozygosity after *X* generations of selfing. Thus, after just four generations of selfing, the average level of homozygosity and the probability that a given locus is homozygous is 94%. However, when *N* loci are considered simultaneously, and if all loci assort independently, then the probability that all *N* loci will be homozygous is equal to $(P_{homo})N$. Thus, if a large number of loci are segregating, then, even after many generations of selfing, there remains a high probability that at least one of those loci remains heterozygous. This prediction has important consequences in the breeding and release of plant varieties that are regarded as homozygous—it means that there are always a few loci that may segregate within the variety. Such phenomena sometimes turn up when a variety is grown in a novel environment where no one has ever tested it before.

Prior to the development of modern breeding methods (and even afterward, for various reasons), plant varieties in self-pollinating species were propagated as mixtures of homozygous lines called landraces. Each landrace typically arose from generations of bulked selections from a farmer's field. Gradually, through selection of desirable types (e.g., large ears of corn) or elimination of undesirable types (e.g., those with seed that fall off during harvest, i.e., shattering), a farmer might develop a particularly useful landrace and share it with friends. Landraces often took on the name of a farmer, a region of origin, a defining characteristic, or a combination thereof (e.g., "Swedish Giant"). It was probably rare for a landrace to originate from a single plant, because there was no knowledge that this might be beneficial, and because this would have required careful multiplication of seed in isolation from other crops before there was adequate seed to plant a crop for harvest. In a landrace that had been grown for many generations, most plants would be homozygous, but the landrace would remain as a heterogeneous mixture of different genotypes.

In 1903, a Danish biologist, Wilhelm Johannsen, reported an important finding that has provided the foundation for modern breeding methods. He showed that progeny grown from a single plant selected from a mixture of inbred lines would produce progeny that were consistently different from those of another plant from the same mixture. Thus, he could create a large-seeded variety and a small-seeded variety through single-plant selections from the same mixture. Importantly, he also observed that further selections within progeny that were derived from the same single plant were not effective. This is because each selection represented a pure homozygous line, and all subsequent variation observed within a selected line was due to differences in environment, and not to genetic differences. A variety selected and multiplied from a single homozygous plant is known as a "pure line," and the alleles or traits possessed by this line are said to be "fixed," meaning that further selection is neither necessary nor possible. These observations, as illustrated in Figure 3.3, are known as Johannsen's pure-line theory. It is also noteworthy that these observations were probably the first time that a clear distinction was made between genotype and phenotype—an important step beyond Mendel's laws.

3.2.4. Other Topics in Population and Quantitative Genetics

Refer back to Figure 3.1, and try to imagine factors that might complicate the situation that is illustrated here. Rather than two genes affecting the trait of melon size, there could be dozens of genes. Different genes could have effects of different magnitudes, and there could also be environmental influences. Rather than being unlinked (assorting independently), some of the loci might be linked together on the same chromosomes. This linkage effect would result in some combinations of parental alleles being more frequent than others. The effects of individual alleles may then not be completely additive as they are in Figure 3.1 (i.e., the genotypic value is found by adding up the individual genotypic effects at each locus). If the genotypic value of the heterozygote is not equal to the average of the homozygotes, we refer to this as *dominance* (meaning that one allele has a dominant effect over another). If alleles at different loci interact (i.e., if the total genotypic value is different from the sum of the genotype values of the individual loci), we call this *epistasis*. Many of these factors will tend to produce a histogram of phenotypes that is more continuous (smoother) than

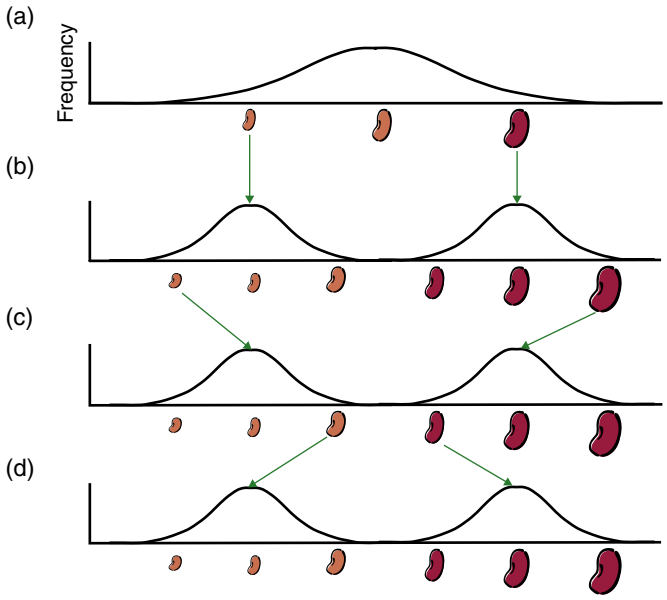


Figure 3.3. Development of pure lines from a mixture of homozygous, heterogeneous beans. Panels (a) through (d) show histograms representing the frequency of different bean sizes in various populations. Panel (a) could represent a landrace or a population derived by repeated selfing of progeny from a hybridized plant. Selection of single beans from the original population results in new populations (b) that have different average bean sizes. Further selection within these populations is not effective (c, d). In this illustration, bean color is a qualitative trait that shows no environmental variation, whereas seed size is a quantitative trait that shows environmental variation. The dark-colored beans on the right represent a pure line, and all phenotypic variation for seed size within that line is environmental. A Danish biologist, Wilhelm Johannsen, conducted a similar series of experiments in 1903, and developed what we now call the *pure-line theory*.

the distribution shown in Figure 3.1, but they can also cause the shape of the distribution to deviate from the normal (bell-shaped) distribution that results when all genes have uniform, additive effects. All of these concepts are simplest to study in a diploid plant. However, many crop plants such as potato and strawberry are not diploids, but rather polyploids. Inheritance in polyploids is considerably more complex than in diploids because there can be more than two different alleles at a given locus, and they can interact in many different ways.

There is an entire field called quantitative genetics that is dedicated to the study and prediction of genetic effects that underlie quantitative traits, and any serious study of plant breeding must be accompanied by further study of quantitative genetics (e.g., see the text by Wrike and Weber (1986)). An excellent introduction to many modern concepts in quantitative genetics is provided by Barton and Keightley (2002). Quantitative genetics builds on the topic of population genetics (the study of gene flow in populations), and many curricula separate these topics into different courses of study.

The study of quantitative genetics has been given a significant boost since the mid-1980s by the discovery of molecular markers, and the ability to produce high-density molecular maps of where these markers and genes lie within plant chromosomes. When mapped molecular markers segregate in the same population as a quantitative trait, it is often possible to find discrete relationships between map locations and individual genes that control the quantitative trait. This procedure, known as quantitative trait locus (QTL) analysis, is the key to understanding the genetic control of many complex traits. It is also the concept that lies at the heart of marker-assisted breeding (the use of molecular markers to assist in the selection of linked traits). A detailed discussion of QTL analysis is provided by Paterson (1998), but an Internet search of "QTL + your favorite plant species" may direct you to primary literature regarding the discovery of QTL in your species of choice.

3.2.5. The Value of a Plant Variety Depends on Many Traits

If melon breeders had nothing to worry about besides fruit size, then melon breeders might have finished their jobs long ago, and/or melons might now be approaching the size of small cars. However, plant varieties are often bought and sold in an open market, and the value of a plant variety is subject to complex and changing industry and consumer preferences. Some of these preferences are mentioned in Section 3.3. There can be no perfect melon variety, but a given market might be driven by the need for melons that are large (but not too large), oblong, sweet, and seedless, and that grow on compact plants that are resistant to insects and disease. The traits that are desirable in crops can change over time. Small families with limited storage might prefer small melons, while those with large extended families might prefer much larger melons. Is there an overall change in the target melon size or must the plant breeder develop both larger and smaller melon varieties? Thus, the perfect plant variety is a distant and moving target, determined by thousands of genes, dozens of which may be segregating in a given population.

3.2.6. A Plant Variety Must Be Environmentally Adapted

Why does a plant variety selected in the tropics not perform well in temperate climates? Many environmental factors influence how a given variety will perform, and genotype–environment interaction (G×E) is an essential concept for breeders to understand. Some environmental factors that interact with genotype include soil type, soil fertility, amount of rainfall, temperature, length of growing season, production methods, and daylength. Some factors such as daylength are predictable, and much is known about how plants respond to daylength. Plants such as soybean require short days to initiate flowering, and there are specific genes that determine when a plant will flower at a given latitude. Other plants, such as oat, require long days to flower, and may flower only in high latitudes during the summer unless specific alleles of a "daylength sensing" gene are absent. However, many other factors that affect G×E are not so well understood. Moreover, many environmental factors such as rainfall are unpredictable, so it is important to select varieties that perform well in a range of environments. This is why plant varieties are tested in numerous locations over a period of at least 2 years before they are sold commercially. If a variety performs well in 1 year at one location, it may perform poorly the next year or at a different location (e.g., Fig. 3.4). Only through multiyear, multilocation testing can we predict how a variety responds to different environments, and whether it will deliver as promised. Related to this is the concept of "stability."

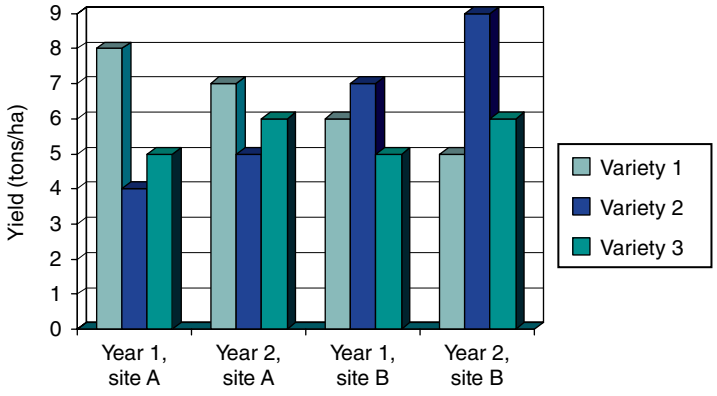


Figure 3.4. An illustration of G×E interaction. Plant variety 1 performs best at site A, whereas variety 2 performs best at site B. Both varieties show variation in performance over years. The fact that the ranking of varieties changes from site to site means that this is a *crossover interaction*. Variety 3 shows performance that is more consistent across environments, so it is described as being more stable than the other two varieties.

If a variety performs consistently in many different environments, we say that it is stable. But a stable variety may not be the top-performing variety in any given environment. Whether to release unstable varieties that perform extremely well in a few environments or stable varieties with average performance in many environments is an ongoing debate among plant breeders. Maize breeders sometimes characterize varieties as racehorses (those that have larger than normal responses to favorable environments), and workhorses (which are more stable but which respond less to favorable environments).

3.2.7. Plant Breeding is a Numbers Game

Before discussing breeding strategies in more detail, it is important to put in context the scale on which plant breeding occurs. As illustrated before, a breeder may be working with multiple objectives and traits in mind that require the selection of many genes. Breeders will seldom know exactly how many important genes are segregating in a population, but there may be information about some of the genes. An example might be a population in which the breeder knows that there are a few specific genes affecting disease resistance, height, and flowering time. Suppose that there are only two genes affecting each of these traits, equaling a total of six genes, all segregating in a population derived from a biparental cross as shown in Figure 3.1. In the F_2 generation, the probability of a specific homozygote at each locus is 1/4. If all six genes assort independently, then the probability of a specific genotype that is homozygous at all six loci is one in 4^6, or 1 in 4096. A breeder who wishes to be reasonably certain of recovering this genotype in the F_2 generation would need to grow many thousands of progeny. Given the fact that many other unknown (or unpredicted) genes will segregate, and that the true genotype is often obscured by the environment, it is not unusual for a breeder to evaluate hundreds or thousands of progeny from a given cross, and to work with many crosses simultaneously. Breeders remark that finding the perfect variety is like winning the lottery. The fact that they often "win something" is a result of "buying many tickets," but the elusive "perfect" jackpot may never be won.

3.2.8. Plant Breeding is an Iterative and Collaborative Process

A common depiction of plant breeding is that it is an ongoing process of gradual improvement, often represented by a gradually upward-sloping graph of yield versus time (e.g., see Fig. 3.5).[3] The sloping line represents the average of many plant varieties released in a given year. The measured performance may be historical, in which case it will reflect changing cultural practice and fluctuation resulting from "good or bad years," or it may be based on a modern experiment in which the performance of older "retired" varieties are tested together with new varieties in the same environment. While the typical graph represents grain yield, many other objectives/traits are selected simultaneously (Section 3.2.5). Therefore, the one-dimensional progress shown in Figure 3.5 does not accurately represent what has been achieved, nor does it account for the fact that objectives and cultural practices change over the years, such that perfection is a moving target.

One might ask "Why not make the perfect cross and select the perfect pure line and be done with it?" The first answer is that the perfect cross cannot possibly contain all the best alleles. Disease resistance may come from one parental source, high protein from another, stem strength from another, and so on. In fact, the perfect parents have probably not been identified for all traits of interest. The perfect variety might require recombining alleles from hundreds of different germplasm sources. The second answer is that the number of progeny required to generate a segregating

[3] An interesting thing about this study, reported by Duvick and Cassman (1999), was that corn varieties showed no improvement if they were grown using the same cultural practices (wide rows) that were used in 1930 to accommodate the driving of horses between rows. Yet, old varieties performed poorly under the modern practice of narrow rows. Thus, the genetic gain that was achieved was accompanied by a trend toward narrow rows: a good example of genotype–environment interaction.

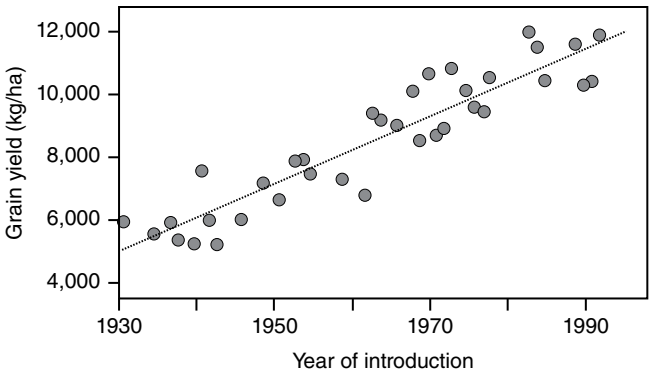


Figure 3.5. Yield of hybrid corn varieties versus year of introduction into agriculture. Data were obtained from Duvick and Cassman (1999), based on field experiments conducted at a plant density of 79,000 plants per hectare at three locations in central Iowa in 1994.

population that contained the perfect combination of alleles would be prohibitive. So the breeder who set out to make a perfect variety would be busy for many decades, while his/her colleagues were busy releasing very good varieties.

Whether working in a cross-pollinating species or a self-pollinating species, the breeder needs to alternate between crossing and selection. Selection is done between crossing generations in order to increase the probability of success, and to release interim varieties. Crossing is done following selection, either to introduce new material, or to recombine existing material. The pedigree of most modern varieties shows a history of crosses that have been made (e.g., see Fig. 3.6). What may not be obvious in Figure 3.6 is that each cross is followed by selection, such that the final outcome is not a random result of the crosses that have been made. Whether it is done intentionally in a systematic process, or ad hoc in an ongoing breeding program, this iterative process of crossing and selection is called recurrent selection. Importantly, plant breeders use material from other breeding programs in their crosses. Legal and ethical principles allow most released plant varieties to be used for crossing purposes in any breeding program, although some commercial plant breeders have restrictions on the use of their varieties as parents, sometimes by using patent protection of their varieties. Furthermore, many breeders actively exchange unreleased germplasm with each other, knowing that reciprocal exchange of germplasm has the net result of increasing the scale and success of each program, thus producing the most public good. Therefore, it is very rare to find a pedigree such as that shown in Figure 3.6 that does not contain material from many different breeding programs, and often from different countries.

3.2.9. Diversity, Adaptation, and Ideotypes

Why does natural or artificial selection not favor a single genotype? Where does genetic variation come from, and why does genetic diversity remain in the presence of intense natural or artificial selection? It is quite clear that genetic diversity originates through mutations in DNA sequence, but when, and how often, do these mutations occur? Why has it been possible, for example, to continually select for higher oil in a population of oats without ever exhausting the genetic diversity? (See Fig. 3.7.) Is it because of new mutations, the unmasking of suppressed genes, the gradual uncovering of rare epistatic gene interactions, the generation of epigenetic changes, or all of these? There have been many interesting debates about these questions, as reviewed by Orr (2005), and literature on this topic makes a fascinating and important side topic for plant breeders.

The famous geneticist, Sewall Wright (1889–1988), introduced the shifting balance theory, in which adaptation and diversity are dependent partially on random population drift (Wright 1982). Although this specific theory is still debated, it is based on two concepts that are highly relevant

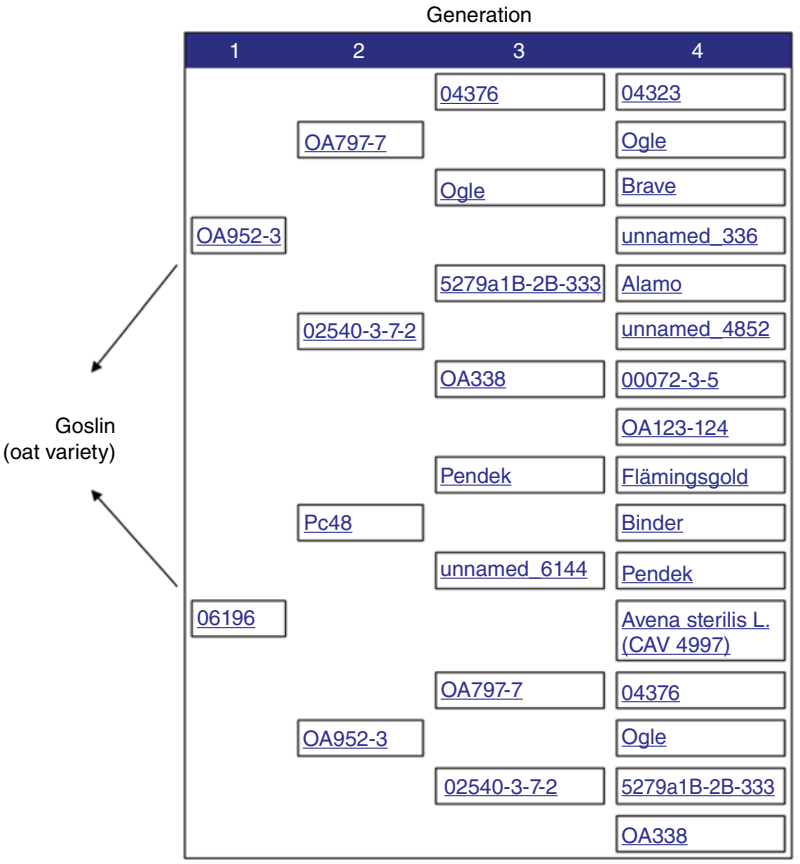


Figure 3.6. The pedigree of an oat variety named "Goslin." The parents of the cross from which Goslin was selected are shown on the left in column 1, grandparents and great-grandparents are shown in columns 2 and 3, and so on. Lines identified by numbers (e.g., OA952-3) were probably elite breeding lines that did not become varieties. This pedigree tree was drawn using an online database (http://pool.aowc.ca/) that records pedigrees of historical oat varieties for many generations.

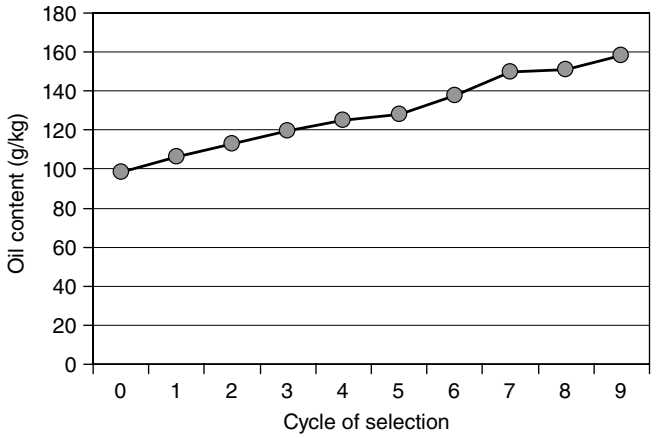


Figure 3.7. Mean oil content for oat lines representing nine cycles of recurrent selection evaluated in three Iowa locations in 1992 (From Frey and Holland (1999)).

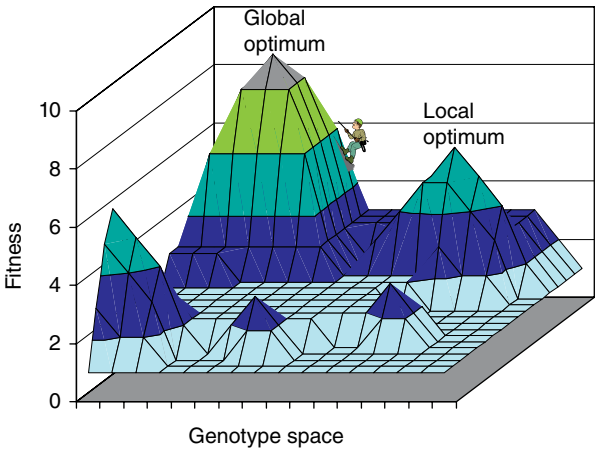


Figure 3.8. In Sewall Wright's shifting balance theory, a genotype or population is defined by coordinates in *N*-dimensional space, and a fitness value forms a surface in the $(N+1)$th-dimension. Here, genotype coordinates are defined in two dimensions on the ground beneath a mountainous fitness surface (the third dimension). The coordinates of a given population can be changed by selection, but only in small increments. Direct selection tends to move a population toward coordinates where fitness is highest, but that may be only a local peak. Applying this concept to plant breeding, we see that to find genotypes beneath a global peak, we need to create and explore a large genotype space (i.e., genetic diversity) or to know exactly where we are going (an ideotype).

to plant breeding: fitness surfaces and adaptive peaks (Fig. 3.8). A fitness surface is a theoretical representation of genotypic value, given an underlying genotype. An adaptive peak represents the genetic coordinates on that surface that produce an optimum phenotype. Adaptive peaks may be local, or global. Part of Wright's shifting balance theory related to how selection was capable of moving a population from one adaptive peak to another, given that selection favors "going uphill." But for a breeder, it is possible to deliberately "go downhill" if it is apparent that this will move a population toward a higher adaptive peak. An example might be the deliberate selection of larger seeds, smaller pods, and a reduced number of pods per plant. Individually, two of these traits would result in poorer yield or adaptation, and the breeder might spend years producing seemingly worthless plants. But once all three traits have been recombined into just the right genotype, the breeder may release a plant variety that achieves a quantum-leap in adaptation. This concept was first formalized under the name "ideotype breeding" by Donald (1968), where an ideotype was defined as a plant with a particular combination of characteristics that have not yet been observed together, but that are predicted to be genetically achievable and are theorized to provide superior yield or adaptation.

The important thing to remember about ideotypes, or fitness surfaces, is that both are dependent on a specific environment. If the environment changes (e.g., if a stress is added or removed), then the value of an ideotype or the shape of a fitness surface could change dramatically. A genotype that was near a global peak on the previous surface may now be in a valley. Given that different agricultural production techniques are really just different environments, it is not difficult to understand why varieties that were adapted to one production practice may not be adapted to another. A variety that is short may produce high yields only when fertilizer is applied, or a variety that does not produce branches may produce high yields only when planted in high-density stands.

By visualizing plant breeding as a mountaineering expedition on a complex and constantly changing fitness surface, you can easily see why plant breeders have vastly different approaches and philosophies. Whether a breeder believes that there is a higher mountain or not, can determine long-term versus short-term success.

3.2.10. Other Considerations

Lack of space in this chapter prevents detailed discussion of many additional topics that make plant breeding challenging and interesting. What follows are some general statements about additional factors that the interested reader may wish to pursue.

- Some crops exhibit *polyploidy* (doubling, tripling, or quadrupling of basic chromosome number). This can lead to nondiploid chromosome pairing, which complicates normal diploid inheritance. Polyploidy can also be induced artificially in order to create "artificial species" that may be more vigorous, or to combine characters from two different species.
- Variation in chromosome numbers (*aneuploidy*) or chromosome structure (*translocations*, *inversions*, *duplications*, and *deletions*) can affect genetic inheritance. These phenomena may also be induced artificially for genetic studies.
- *Xenia*, the expression of pollen genes in the tetraploid endosperm or embryo of a seed, can complicate the selection of seed traits. Seeds normally show the phenotype of the parent plant, but carry the genotype of a genetically different offspring.
- Breeders are important stakeholders in efforts to maintain biodiversity through *in situ* and *ex situ* collections of germplasm.
- Interspecific hybridization has been used to transfer alleles controlling traits such as disease resistance, which may not be present in the normal germplasm of a species.
- Male sterility, and various methods of pollen control, may be useful in the production of hybrid varieties, but have also been used in normal recurrent crossing programs.
- Plant breeders must manage large amounts of data, and they may need to share data with other researchers. Electronic data management systems are becoming increasingly important in plant breeding.
- Plant breeders must also be statisticians. There is a large body of literature concerning the optimization of field plot techniques, and the statistical analysis of test results.
- Resistance to disease is an ongoing battle between plant breeders and the organisms that cause disease (a variant of the "evolutionary arms race"). Many pest organisms mutate very quickly, and mutations that overcome new types of resistance are selected quickly in crop monocultures of a single plant variety. One strategy to overcome this is the development of *multiline varieties*, which contain a mixture of resistance types. Another method is to "pyramid" or "stack" multiple sources of resistance into a single variety.

3.3. OBJECTIVES IN PLANT BREEDING

The overall value of a plant variety is determined by many small and subtle characteristics that are quantitative and polygenic in addition to a few major characteristics that may be qualitative and monogenic. It is not difficult to draw a parallel in human traits. Think of someone whom you admire; you might like this person because you have a fondness for a certain eye color, hair color, height, or other characteristic, but most likely it would be because of a combination of many subtle traits (involving both appearance and personality) that are controlled by numerous genes but also influenced by environment. Beauty is in the eye of the beholder, or in this case, in the eye of the plant breeder.

Many breeding objectives fall into two general categories of traits. We often categorize certain traits as agronomic traits or input traits because they relate to production practices and to the amount of raw material that can be harvested. Such traits include crop yield, pest resistance, height, flowering time, susceptibility to lodging (falling down), seed vigor, and seed dormancy. Crop yield includes many component traits, such as seed size, seeds per pod, pods per branch, and branches per plant. Some breeders prefer to select according to component traits rather than on final yield, but the value

of a plant variety is almost always judged for its potential to produce high yields per unit area. The second general category is described as output traits, which include anything related to the composition or quality of what gets harvested. Examples are the composition and content of protein or oil; the relative proportions of oil, starch, and protein; and the composition of secondary compounds that may have value relative to human health and industrial use. Many output traits are extremely complex. These include traits related to the use of a plant product in processing.

For every cultivated plant species, there are different breeding objectives. Often, there are several different sets of objectives, sometimes conflicting, for the same species. An example of this is a barley variety, which may be used for animal feed or malt production for beer. A major objective for animal feed is high protein content, whereas malt production requires low protein content. The breeding objectives for crops such as malting barley or bread wheat are dependent on markets that have evolved very specific industrial processes that require dependable and uniform grain. Thus, a variety that is merely high-yielding and pest-resistant will not suffice. Other crops have highly diverse objectives that may be driven by many different markets. Soybean, for example, is used in the manufacture of many different food products, and each product has very specific requirements such as taste, texture, chemical composition, seed size, and seed coat color.

Horticultural crops can have many interesting and diverse objectives. The market for gardening varieties is driven by diversity because gardeners like to try new things. In contrast, the market for commercial varieties of horticultural crops is driven by the need for conformity and uniformity, and for improvement of specific high-value characteristics. For example, processing tomatoes need to deliver maximum amounts of soluble solids, but there would be likely be little tolerance for a variety that had a non-tomato taste or different color. The market for fresh shipping tomatoes has requirements for produce that is both durable in transit and attractive in appearance (and, some would argue, tasteless).

Some breeding objectives can produce interesting challenges. Consider the objective of reduced seed content in grapes or watermelons. Where does the seed come from to grow the next crop? In these cases, breeders can "trick" the plant by producing hybrid varieties that have an unbalanced number of chromosomes such that they cannot undergo proper meiosis to produce viable seed. Another interesting challenge is the incorporation of traits that we cannot measure directly. For example, it is desirable to incorporate multiple sources of disease resistance in a single plant variety so that it is more difficult for a pathogen to mutate and overcome the resistance. But if there are two genes at different loci that both confer resistance, how do you know that they are both there? In this case, one solution is to identify genetic (DNA) markers that are linked to each resistance gene, so that the markers can be selected instead of the resistance.

Breeding objectives can also change suddenly and unpredictably. An example was the sudden appearance in the 1990s of devastating levels of a fungal pathogen called Fusarium in wheat. These days every major wheat breeding program focuses on introducing new sources of resistance to Fusarium disease. Another example was the development in the United States of a health claim for soluble β-glucan fiber in oat. While breeders were aware of this factor, and had initiated selections for higher β-glucan, it was not until the development of this health claim that a particular value was placed on varieties that had elevated levels of this trait. These points illustrate that breeders must be constantly aware of changing market forces and agricultural conditions. Indeed, it is not unusual for breeders to help drive some market forces through their knowledge of traits that are available, and through awareness of economic factors affecting industry and producers.

3.4. METHODS OF PLANT BREEDING

Typically, breeding is a continuous cycle of recurrent mating and selection. There is rarely a start-point or an endpoint in a breeding program; rather, it is a continuous pipeline that must be kept filled for continual delivery of new and better plant varieties. Breeders try to release improved varieties

every year, but today's varieties may be the result of planning and crossing that began a decade ago. Add to this the fact that breeders mix and match various breeding methods, depending on objectives, and that they constantly modify and update their strategies, it is understandable why it is difficult to write down a simple "recipe" for successful plant breeding. Nevertheless, several core strategies have been developed, and most breeders adopt and adapt one or more of these strategies depending on plant characteristics, breeding objectives, resources, and personal preference. While the breeding systems that are described in the text appear to have a beginning and an end, many cycles at different stages will be running simultaneously in a given breeding program, and that material at the end of one system can become starting material for another system.

3.4.1. Methods of Hybridization

Most breeding methods incorporate sexual hybridization as a method of generating new genetic variability. Hybridization may occur naturally, as in the case of out-crossing species, or it may require tedious manipulation of flowers in the case of a self-pollinating species. In special cases, sexual hybridization has been used as a method to combine traits from species that are rarely cross-fertile. The methods for hybridizing most self-pollinating species involve emasculation (removal of stamens) and the introduction of pollen from another plant. The timing of these steps, and the methods by which they are best done, are critical. Outcrossing species may also require controlled hybridization in specific breeding methods, particularly if hybrid varieties are developed (see Section 3.4.3.4). Even varieties that are developed through random mating require special considerations. For example, alfalfa is poorly pollinated by honeybees, which do not trigger a special floral mechanism that transfers pollen to the bee, but they are efficiently pollinated by certain wild bees, which may be artificially reared near plots that are used for breeding or seed production. Fehr and Hadley (1980) have compiled a comprehensive reference source on methods of hybridization in crop plants that discusses technical details as well as many related issues such as environmental factors that affect the timing of flowering and fertilization.

3.4.2. Self-Pollinated Species

Most self-pollinated species are grown as varieties derived from a pure line (see Section 3.2.3). Therefore, the overall objective of the following strategies is to recombine as many desirable genes as possible into a single homozygous genotype. All of the following strategies involve one or more hybridizations followed by generations of selfing and selection. The key differences among these strategies are whether crossing is repeated, when selections are made, and how many selfed progeny are made from each selected plant. All systems generally culminate in the same final steps for variety testing and release.

3.4.2.1. Pedigree Breeding. The pedigree breeding method (Fig. 3.9) requires detailed record keeping. Selections are made in every generation except for the F_1 because it is assumed that all F_1 plants from a cross are genetically identical. This method is chosen because it allows the elimination of poor lines, or selection of traits with high heritability at an early stage in the breeding program, thus leaving more room to increase the number of lines that can be tested from promising families. An additional benefit is that, by recording information about the performance of lines as well as their parents and families, the breeder ensures that selections can incorporate all three types of information. For example, one family might appear to be susceptibility to disease, while another family from the same cross might appear to be completely disease-resistant. This might lead to speculation that the first family was derived from a parent that was segregating for disease resistance, while the second family was derived from a parent where the resistance was fixed. This is useful information, since individual lines sometimes escape disease infection even if they do not carry genetic resistance. This information might allow the breeder to favor selections within the resistant family.

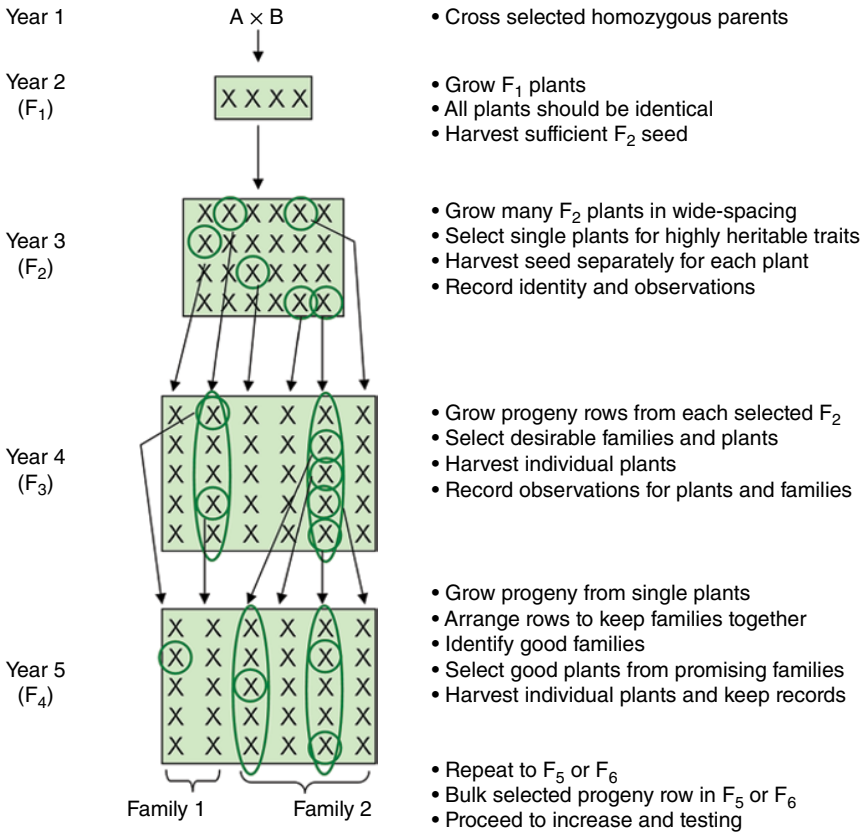


Figure 3.9 The pedigree breeding method is used in self-pollinated species to derive pure-line varieties when it is desirable to practice selection in early generations.

3.4.2.2. Single-Seed Descent. The pedigree method has two downsides. First, it requires a lot of time and resources to keep records about material that will simply be discarded. Second, performance of progeny in early generations may be enhanced by the effects of dominance, which is lost in later generations, and also that favorable gene interactions (epitasis) may not be evident until later generations. In other words, a good line in an early generation may give poor progeny in late generations, or a poor line in an early generation may give good progeny in late generations.

The SSD method (Fig. 3.10) addresses all the potential downsides of the pedigree method. Rather than select lines and families in early generations, a large F_2 population is created, and one random line is developed from each F_2 plant. Thus, the pedigree of each F_2 line is represented by exactly one random line in each following generation by taking a single seed from each F_2 family in every segregating generation. The result is that maximum genetic diversity is preserved until late generations when selection will commence, and no recordkeeping is required. The SSD method has the advantage that off-season or winter nurseries (e.g., in Puerto Rico or Argentina for North American crops) can be used to get an additional one or two generations per year. Since only one or a few seeds are needed from each plant to advance to the next generation, optimal plant growth is not required in the winter nursery. Rather, quick seed maturity is favored in winter nurseries for SSD breeding. The SSD method can also be used to derive populations of recombinant inbred lines (RILs). These populations are useful in genetic experiments or QTL discovery because segregation can be considered random such that "good" lines can be contrasted with "bad" lines to identify genetic determinants. However, it is this same feature that leads to the primary criticism of SSD as a breeding method: poor material that might easily be removed in early generations continues to occupy space and resources in the breeding program.

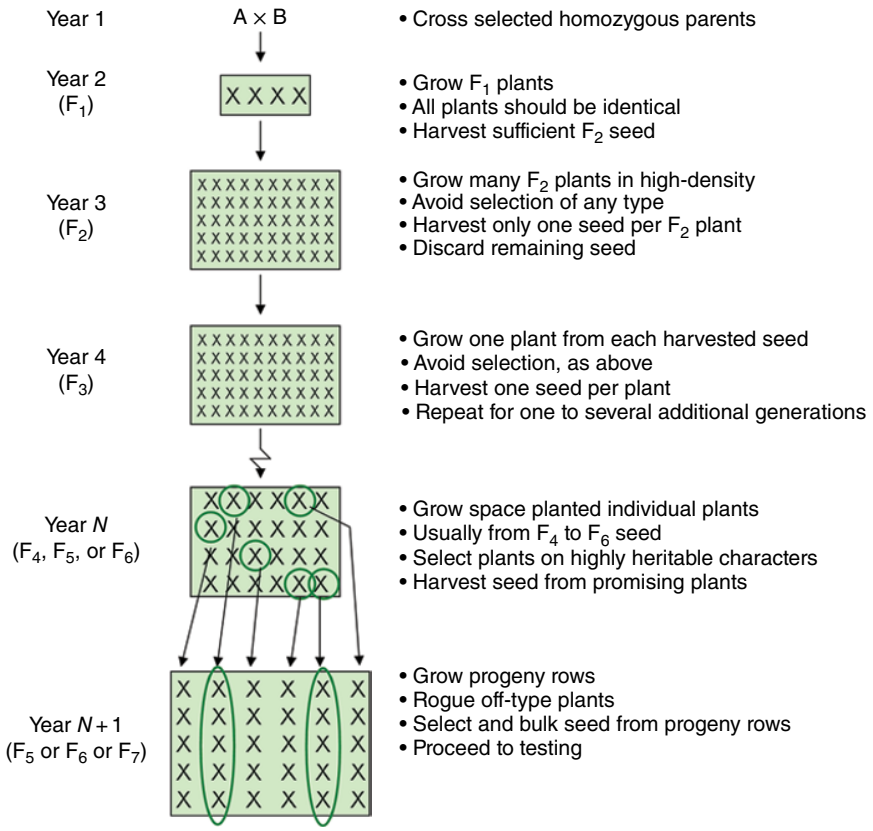


Figure 3.10 The single-seed descent (SSD) breeding method is used in self-pollinated species to derive pure-line varieties when it is desirable to select from random homozygous lines in an advanced generation.

3.4.2.3. Bulk Breeding Methods. A bulk breeding method is any method whereby generations are advanced by bulking and planting seed from the previous generation. However, if all seed from a given generation is harvested, then there will likely be too much seed to plant in the following generation, so some seed must be discarded or held in reserve. The SSD method is actually a special type of bulk breeding whereby each generation is advanced by saving only one seed per plant from the previous generation. In other methods of bulk breeding, the reduction of seed is achieved randomly, or through a selection process that is applied uniformly but indiscriminately (e.g., harvesting the earliest-ripening plants, or sieving to keep the largest seed). Bulk breeding is often a desired method of generation advance because it is extremely simple. However, several issues must be considered. If the intention is to preserve a random nonselected population, then some lines will be represented by greater numbers of progeny than other lines simply by random chance; thus, the total genetic variation that is preserved is less than that of a true SSD population. Lines that are preserved are unlikely to be random. Plants that have more seed will be disproportionately represented, and plants that compete poorly may be lost completely from the lineage. This may sound like a desirable way to favor lines that are more adapted. The evolutionary plant breeding method uses natural selection during bulk generation advance to provide selection pressure. Other breeders practice bulk generation advance in the presence of some artificial mass selection. Examples include favoring tall plants by mechanically harvesting only the tops of the tallest plants, or conversely, penalizing tall plants by applying herbicide using a rope-wick prior to harvest. Many creative methods have been developed to apply mass selection during bulk generation advance. However, it must be remembered that plants that produce more seed during generation advance might actually produce less seed in a

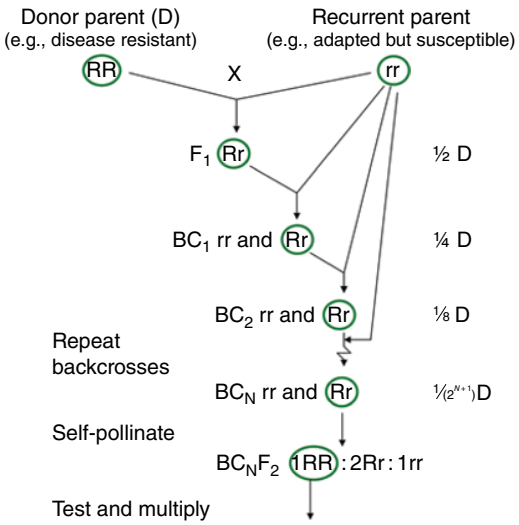


Figure 3.11. The backcross breeding method is used to transfer alleles at a small number of loci from a donor parent into the genetic background of a reciprocal parent. Each generation of backcrossing reduces the proportion of alleles from the donor (D) parent by half (1/2), as shown on the right. Note the similarity to this method and what can occur naturally (Fig. 2.13).

competitive community of identical genotypes, or they may simply produce seed that is smaller and less desirable. In crops where seed yield and late maturity are associated, the plants that produce more seed are later maturing and the bulk population will have its maturity shifted later.

3.4.2.4. Backcross Breeding. As illustrated in Figure 3.11, the backcross breeding method is quite different from the methods discussed so far. It involves much smaller populations and greater numbers of hybridizations. The objective of a backcrossing program is to preserve as many genes as possible in an inbred parent that has proven adaptation to a given environment, while introducing new alleles at just one or a few loci from an unadapted parent. The former is called the recurrent parent, and the latter is called the donor parent. Often, a backcross strategy is used when an unadapted genotype is found to have disease resistance that is not present in adapted varieties, but it may be used to introduce any simply inherited trait from a parent that is otherwise undesirable. This method is also very useful for moving a transgene or foreign DNA introduced via biotechnology into the recurrent line, and also can occur in nature (see Fig. 2.13).

After an initial cross between the donor and recurrent parent, repeated backcrosses are made to the recurrent parent. Each time a backcross is made, the progeny receive half their alleles from each parent; thus, the proportion of alleles that remain from the donor parent are reduced by 50% each generation, and after N backcrosses, the proportion of alleles remaining from the donor is formulated as $1/(2^{N+1})$.[4]

After the first hybridization, every backcross will produce a mixture of genotypes at the locus where we wish to introduce alleles from the donor. Thus, each generation must produce an adequate number of progeny such that it is possible to identify the heterozygous genotype carrying the donor allele. While Figure 3.11 shows an example of backcrossing alleles at a single locus, it is also possible to backcross alleles at two or more loci, but larger numbers of progeny must be made to identify backcross parents that contain donor alleles at all loci.

[4]This is true only for loci that are not linked to the locus under selection. Donor parent alleles that are closely linked to the locus under selection will be highly favored. This phenomenon is called "linkage drag."

3.4.2.5. Testing Pure-Line Varieties. At the end of these inbreeding procedures (typically at the F_5, F_6, or BC_5 stage), most breeders will consider that individual plants are adequately homozygous to form a pure line that could be tested for lower heritability traits, such as seed yield. Additional selection among inbred lines must be performed. This selection will be among a decreasing number of lines that are increasingly elite performers. This selection may result in the complete elimination of all lines from a given cross—perhaps in favor of keeping additional lines from a cross that turned out more favorably. The process of testing begins with discarding obviously inferior lines and ends with selecting the few best lines each year. For seed crops, selection for high seed yield is the primary objective of the testing program. Selection begins with many lines tested at one or two locations with little-to-no replication. As inferior lines are discarded and fewer lines remain in the program, testing moves to multiple locations with more replication. In this progression toward more intensive testing, selection becomes more accurate for traits with lower heritability. As superior lines are identified for commercialization, they will need to enter a seed certification system.

3.4.2.6. Seed Increase and Certified Seed Production. As superior lines are identified, preparations for seed certification and seed increase will begin. An inbred line may be required to meet established standards for registration as an "official" plant variety. These standards include a written description for crop inspectors, uniformity that ensures the variety is similar from year to year, uniqueness that ensures a variety is different from other varieties, and sometimes merit in relation to competing varieties. Increasingly, plant variety protection or "plant breeders' rights," which restricts unauthorized use of the variety, is sought in addition (or as an alternative) to variety registration.

Seed certification schemes may exist at the state, national, or international levels. Seed certification is a guarantee that the grower receives seed from the stated variety. Seed certification entails a description of the variety, crop and seed inspections, and the number of generations that may be used to multiply seed for sale to growers.

These requirements for variety registration and certification may vary among states, nations or international levels, and they are often unique among crop species. However, the general principles are the same, and so is that fact that they are often addressed simultaneously. Table 3.1 lists the typical final steps in the birth of a plant variety, and how each of the abovementioned requirements might be met.

TABLE 3.1. Steps Involved in the Final Stages of Variety Development in a Self-pollinated Species

Step	Description	Activities
Progeny rows	Single row, nonreplicated	Select promising rows, rogue off-types, bulk seed from characteristic plants
Home tests	One site, two replicates, 1 year: material from home program + standard check variety	Detailed selection may include characters such as quality that require increased amounts of seed; rogue off-types and bulk remaining seed
Preliminary tests	Three sites, three replicates, 1 year: material from home program + standard check variety + other current varieties and possibly lines from collaborating breeders	Detailed and final selection of several lines to enter into variety registration trials; identify defining characters, rogue off-types, bulk seed
Breeders' seed rows	Nonreplicated plots grown in isolation to increase seed for supply of potential variety	Identify defining characters, rogue off-types, bulk seed
Registration tests	Six sites, four replicates, 2 years: cooperative tests that include several best lines from each breeder + standard check varieties	Selection of lines to support for registration (often competitive among breeders); apply for plant variety protection on probable winners (year 2)

Note: These steps must address several requirements simultaneously: testing and selection, varietal purity, registration and/or protection, and seed increase and distribution.

Typically, the steps listed in Table 3.1 technically begin after the last generation in which a family can be traced to the progeny from a single plant (the founding generation). If that were the F_5 generation, then the resulting variety would be F_5-derived. Since homozygosity continues to increase with each generation of selfing, breeders often use the notation $F_{X:Y}$, where X denotes the founding generation and Y designates generations of selfing and bulking that follow the founding generation. Varieties developed from early founding generations (i.e., a relatively heterozygous founding plant) can show a considerable amount of heterogeneity, especially in molecular traits that do not undergo selection. Rogueing (the culling of undesirable plants or off-types) is done in generations following the founding generation, and this reduces the amount of heterogeneity. However, it is common for phenotypic variation to show up within a variety, sometimes by surprise, when a new environment is encountered. For this reason, varieties may need to be described in terms of their range in characteristics, and descriptions based on molecular traits are increasingly favored.

Although they are not technically part of a breeding program, states (and other jurisdictions) have variety testing programs. These programs are usually run by government, university, or producer associations to enable plant varieties to be compared under different conditions (years, locations, and management practices). The resulting data allow for recommendations to growers who can choose varieties best-adapted to the growers' conditions. If all works well, the entire process leads to a productive crop for the grower.

3.4.3. Outcrossing Species

Since pure lines cannot be easily maintained in a naturally outcrossing species, the development and release of varieties in an outcrossing species is quite different from that in a self-pollinated species. Rather than identifying the "perfect genotype," the objective is to identify the perfect set of genes that work happily together in a random mating population. Some outcrossing plant species, such as rye, can tolerate a high degree of inbreeding, and can be effectively bred and grown as if they were self-pollinated species. The primary difference is the increased need for isolation (to prevent uncontrolled outcrossing) during variety development and seed production. However, other species do not tolerate inbreeding, or they have specific mechanisms to prevent it. Matings between different plants often produce offspring that are more fit than the parents: a concept called hybrid vigor or heterosis. Thus, maintaining a heterogeneous population in a random mating state is beneficial. However, there must still be some opportunity to select the breeding population such that it produces relatively uniform progeny that have desired and predictable characteristics.

3.4.3.1. Mass Selection. Historically, varieties of outcrossing species were improved in the same way that landraces of self-pollinated species were improved: by saving seed in bulk, or by saving seed from selected plants. Either way, selection would have taken place. For example, we know that selection for seeds that do not fall off the plant before manual harvest (called "shattering") was one of the earliest traits selected in the process of crop domestication. This most likely happened simply because genotypes with seeds that shattered were rapidly eliminated as soon as early agronomists started planting crops intentionally from seed that was harvested for food rather than gathering seeds from plants in the wild. Other traits, such as lack of seed dormancy, seed size, early flowering, and height, were probably selected in similar ways, whereas traits such as fruit flavor may have required a more deliberate effort to propagate favored genotypes. Regardless of whether selection is deliberate or not, all strategies that allow a limited part of a population to advance to a new generation are loosely termed *mass selection* (Fig. 3.12). The primary difference between mass selection in a traditional open-pollinated variety and that in a traditional landrace is that the former undergoes continual random mating. This difference is the reason why mass selection has fallen from favor in self-pollinated crops, where continual recombination does not take place and mass selection is only used during the first few generations, following a cross when there is variability within which to select. Mass selection is still a viable method to improve a cross-pollinated species. With modern knowledge

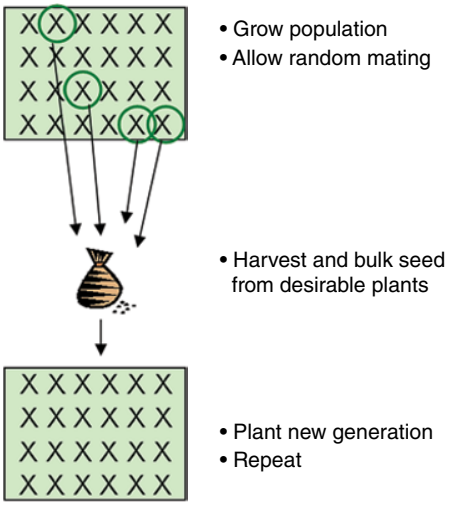


Figure 3.12 Mass selection, as practiced in an outcrossing species, is a traditional method of breeding that is still used to improve base populations from which parents may be chosen for other breeding methods.

about genetic diversity, a mass selection strategy will now try to reconcile the intensity of selection with the need to maintain diversity. Mass selection is often used as a strategy for continual population improvement in crops such as maize, although it is now more likely to be used to improve a base population that will serve as a source of germplasm for other breeding strategies. Mass selection can also allow introduction of new or exotic germplasm that will recombine with an elite population.

3.4.3.2. Recurrent Selection. The term "recurrent selection" has been used earlier in this chapter to refer to any strategy where selection is alternated with recombination (Section 3.2.8). In fact, mass selection is technically a type of recurrent selection, because recombination occurs with every generation. However, special recurrent selection strategies have been devised for cross-pollinated species whereby selection and intermating are more discrete, and controlled.

An important modification over mass selection has been the development of methods whereby plant selection is based on the performance of their progeny. This is highly relevant if one wishes to favor genes that increase the fitness of the population. Mass selection merely saves plants that have a desirable phenotype, but there is no guarantee that the alleles controlling this phenotype will be expressed in the same way when they are mixed with other alleles in the following generation. Figure 3.13 shows a recurrent selection strategy that allows full progeny testing prior to random mating. It is noteworthy that this strategy is based on a cycle that requires multiple years to complete. Other methods of recurrent selection have been devised, some that require an additional generation of pollen control, and others where some level of controlled mating and selection can be performed every year. Important considerations in the selection of a recurrent selection strategy are the heritability of the trait, and the time at which it is expressed. Traits with low heritability, such as grain yield, are more responsive to progeny testing. It is very difficult to predict crop yield from a single plant, so it is far more accurate to test yield in a whole row of progeny. Traits with high heritability can be selected on a single-plant basis, and traits that are expressed prior to pollination can be selected more effectively by eliminating those that do not express the desired phenotype prior to pollination. Recurrent selection is often used to develop base populations from which other forms of selection or crop improvement can be made.

3.4.3.3. Synthetic Varieties. Most modern agricultural practices require plant varieties that are predictable and uniform. Even though mass- and recurrent selection are practiced to improve base

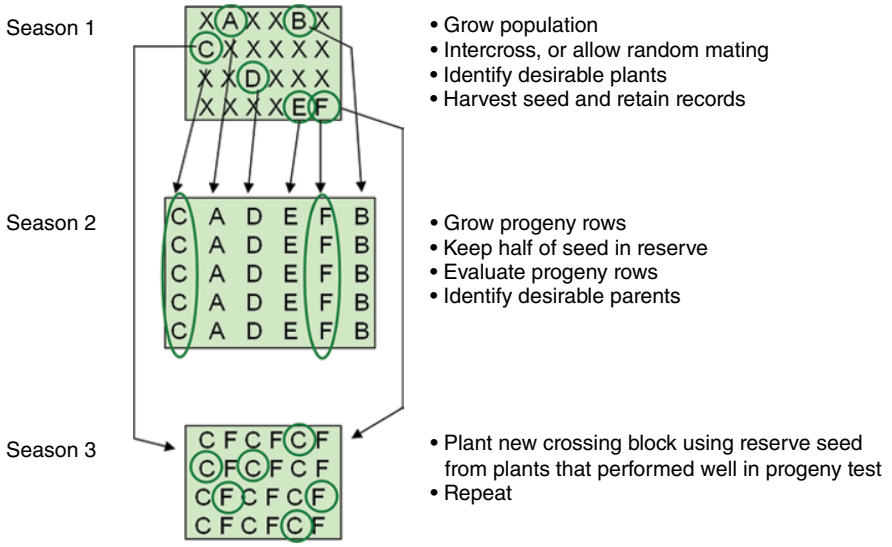


Figure 3.13. An example of a recurrent selection strategy with progeny testing. Many variations on this type of strategy have been devised.

populations, these populations may be too variable for modern production practices, or they may be difficult to maintain in a state that will perform as predicted. A mass-selected population may continue to improve with time, but it might also be inadvertently selected into a state that could theoretically cause damage or liability. Consider an alfalfa variety that has been selected to remain in the vegetative state for an extended period of time. This trait might be a desirable characteristic if the crop is used for repeated cutting to produce green forage for animal feed, but not helpful to produce seed for breeding or sales.

A synthetic variety is an early random mating population derived from a mixture of a group of "reproducible components" (Fig. 3.14). The components can be inbred lines, clones, or hybrids. For example, in perennial forages, synthetics are initiated from a small set of parental lines with proven merit in progeny tests. Because they are perennial, these parents can be maintained indefinitely, and can be propagated vegetatively in small quantities. It is also advantageous that these species produce large quantities of small seed. In other species, a synthetic variety may be initiated from inbred lines. Equal quantities of intercrossed seed from each founding line (Syn-0) is harvested and used to plant a progeny generation called Syn-1. Seed from Syn-1 may be sold as a variety or used to produce a next generation called Syn-2. Generally, having few intercrossing generations is desirable, but the number of generations will be determined primarily by limitations of the species and requirement for seed. In order to maintain uniformity and vigor, synthetic varieties must be reconstituted regularly.

3.4.3.4. Hybrid Varieties. A hybrid variety is a special type of synthetic variety that is defined as the first or second generation derived from crosses among inbred lines. Historically, many hybrid varieties were composed of double crosses (e.g., [A × B] × [C × D]) or three-way crosses ([A × B] × C). However, most modern hybrids are now produced from seed that results directly from the mating of two inbred lines (Fig. 3.15). Seed from such a mating is expected to be highly heterozygous and highly homogeneous: two attractive traits for most crops. The primary advantage of a hybrid variety is that it can provide a performance advantage resulting from heterotic effects at many heterozygous loci, yet be highly uniform and predictable. A disadvantage is that the uniformity lasts for only one generation, so seed must always be purchased from a hybrid seed production facility. This is an advantage if you are the hybrid seed producer, or if you are a plant breeding company that needs to

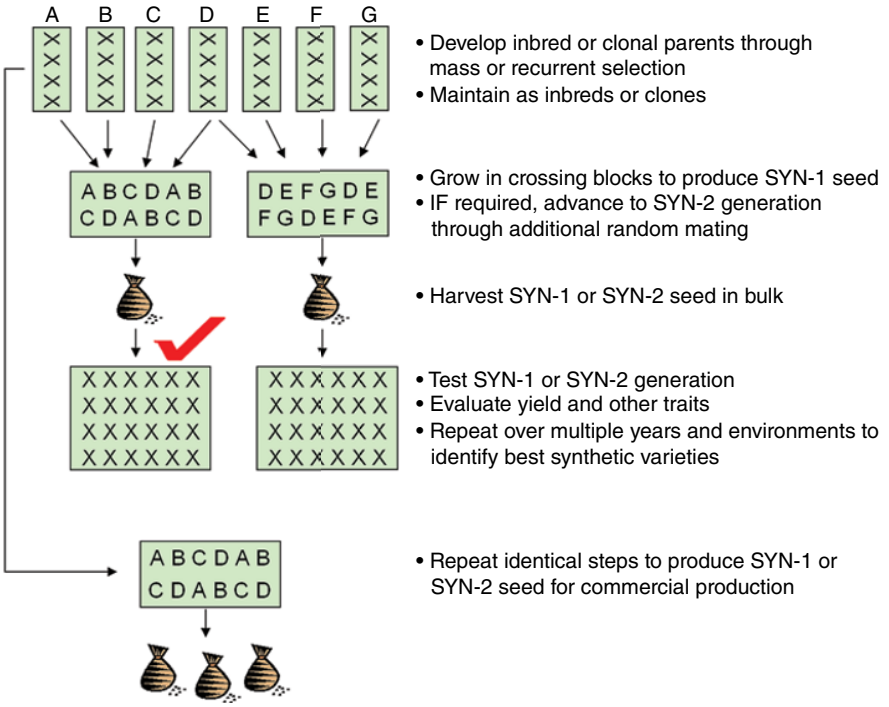


Figure 3.14. Schematic simplification of the development of a synthetic plant variety in an outcrossing species. The Syn-1 generation is produced by random mating of reproducible components (inbred lines or clones). If it is found to be desirable as a new plant variety, it can be reproduced and sold by repeating the identical crossing block. This type of breeding method is most practical in a perennial forage species. If adequate seed cannot be produced in Syn-1 generation, the Syn-2 generation (harvested from Syn-1) may be used instead.

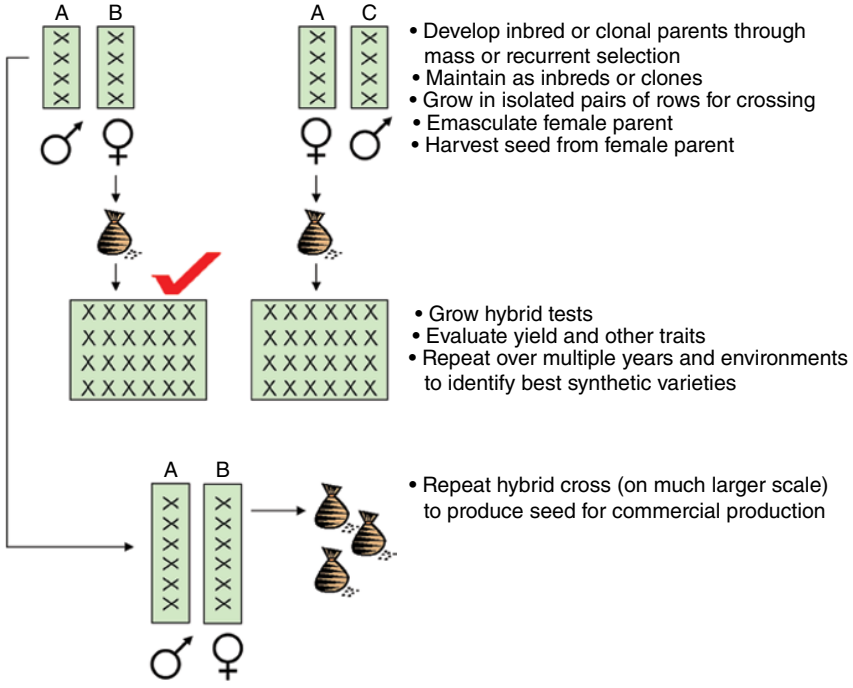


Figure 3.15. Schematic simplification of the development of a hybrid plant variety. In corn, the parents (i.e., A, B, and C) are inbred lines that have been derived through other breeding methods. In other crops, the parents may be clonally propagated. Parents are grown in adjacent rows for crossing, and the female parent is emasculated so that it will not self-pollinate. Seed harvested from the female parent is tested in performance trials. If a hybrid variety is successful, the cross is repeated on a large scale for commercial production.

control the distribution of your variety. Farmers are willing to pay the added cost in many cases because the hybrid varieties have uniformly good properties that lead to assured production and high yield.

Maize is a crop that is normally cross-pollinated, but it can easily be self-pollinated in order to derive inbred lines. For this reason, and because maize shows a large amount of hybrid vigor, most modern maize production is based on hybrid varieties. Because of the high value of the maize industry in many developed countries, the development of hybrid varieties, and subsequent hybrid seed production, is dominated by industry.

Hybrid variety development involves extensive testing of many different hybrid crosses that are developed by intercrossing inbred lines in many different combinations. Inbred line development is not random. Inbred lines are generally developed in two streams, called heterotic groups, from two different genetic backgrounds that are known to produce good hybrids when crossed with each other. Lines in each stream are frequently test-crossed with each other using elaborate schemes that can help guide inbred line selection in both streams. New or unknown inbred lines are typically test crossed to both heterotic groups to determine which test cross shows the most heterosis so that the new line can be placed into the appropriate group. Inbred lines, and the populations from which they are derived, are carefully guarded secrets of every commercial maize breeding company.

Hybrid maize seed production involves planting alternate rows of two different inbred lines, and removing the male parts (tassels) of the line intended for hybrid seed harvest, that is, the "female" parent. The removal of tassels must be done carefully before they emerge and shed pollen.

Hybrid varieties can theoretically be produced in any crop species; but for some species, it is not practical or commercially viable. In many cross-pollinated species, inbred lines are difficult or impossible to produce, and so synthetic varieties are used. In self-pollinated varieties, hybrids may not show as much advantage as they do in cross-pollinated varieties, because these species naturally tolerate inbreeding. Furthermore, it is more difficult to manually enforce an adequate number of hybrid matings. However, hybrid varieties are frequently used in high-value horticultural crops that produce a large amount of seed from a single mating. Hybrids have also been used in some self-pollinated crops in which mechanisms of male fertility can be used to ensure cross-fertilization. Some of these crops include sugar beet, sunflower, and rice.

3.4.4. Clonally Propagated Species

Some crop plants are propagated naturally and/or artificially through vegetative propagation rather than through sexually produced seeds. Globally, the most important example is potato, but other examples include banana, strawberry, yam, sweet potato, and many tree crops. Although the crossing behavior of clonal crops is not relevant to propagation, it is still important in the breeding strategy. Most clonally propagated crops are cross-pollinated, so breeding methods are most similar to those used in cross-pollinating seed crops. However, the ability to maintain an "immortal" genotype makes selection of a population less important, and selection of individual plants becomes far more relevant. Most of the bananas in commercial trade are a single variety. In the 1950s, the single variety of the day was "Gros Michel," which was wiped out by Panama disease. A new variety, 'Cavendish,' was selected since it was resistant to Panama disease; but since the 1990s, a new Panama disease strain is now attacking 'Cavendish' around the banana cultivation world and is a threat. While wide adoptation of a single genotype is possible with clonal propagation, this practice leaves production vulnerable to a new disease strain and other pests. The selection of tree crops presents special challenges because of the long juvenile period, so many fruit tree varieties have been identified by careful observation of hybrids from serendipitous crosses that may have taken place many years ago.

3.5. BREEDING ENHANCEMENTS

This section provides a brief description of several of the most important techniques that can be used to enhance the success of plant breeding. Many additional techniques are discussed in other literature. Perhaps, the most important modern breeding enhancement is plant transformation: the ability to transform plants with DNA that originates from different species. This topic is discussed in other chapters, but it is interesting to note that the way in which genetic transformation can be incorporated into a breeding program bears many resemblances to the use of mutation breeding, discussed in Section 3.5.3. Furthermore, marker-assisted selection or MAS (Section 3.5.2) is often used as a follow-up to genetic transformation in order to recombine a transformation event into new breeding populations.

3.5.1. Doubled Haploidy

The derivation of genetically pure lines (Fig. 3.3) is one of the most important steps in breeding self-pollinated varieties. In the SSD method (Fig. 3.10), this step could be considered "wasted time" if there were a shortcut to produce pure lines. This shortcut exists, and it is called doubled haploidy.

The principle behind doubled haploidy is that every plant species produces haploid gametes during meiosis. Haploid gametes are found in the female (egg) and in the male (pollen) tissues. By forcing these gametes to double the chromosomes in their nuclei, we can immediately produce a cell type that is both diploid and homozygous. There are several techniques by which this phenomenon can be induced. The most common technique is through artificial culture of the male gamete (microspore) or the tissue containing those gametes (anthers). By culturing those tissues on a growth medium, we cause the haploid cells to undergo mitotic divisions. At some stage, a natural doubling may take place when mitotic nuclei fail to divide into separate cells. Doubling can also be induced by the addition of a chemical called colchicine that interferes with normal cell division. The culture can then be forced to regenerate into a normal diploid plant (see Chapter 5).

Techniques to produce doubled haploids have been developed in many crop species, and the technique is used routinely in species such as wheat, barley, and canola. While the cost of producing doubled haploids is often greater than that of producing SSD lines, the ability to accelerate the breeding program (by skipping three or four self-pollinating generations) and get to market faster with a new variety is often worth the added cost.

3.5.2. Marker-Assisted Selection

Earlier it was mentioned that individual genes contributing to complex plant traits can sometimes be discovered through their association with genetic markers. This procedure, called QTL analysis, provides the foundation for a more efficient type of genetic selection called MAS (Fig. 3.16). Rather than selecting traits, which are the outcome of many genes, MAS is based on selecting specific alleles at marker loci that are known to be linked to the genes that cause the desired trait. The theoretical advantages of MAS are that it (a) avoids errors caused by environmental variance; (b) can be applied at a juvenile stage before a trait is expressed; (c) can be applied on a single plant, whereas phenotypic selection of some traits might require seed or tissue from many plants to be effective; and (d) may be less expensive than phenotypic selection. Although MAS does not replace the requirement for parent selection, sexual recombination, and breeding strategies, it can significantly increase the efficiency by which superior genotypes are selected. For this reason, MAS is considered to be an important modern enhancement of traditional plant breeding.

The theoretical advantages of MAS may not always be relevant, and it is often argued that phenotypic selection is faster and cheaper than MAS for many traits. Some of the factors that

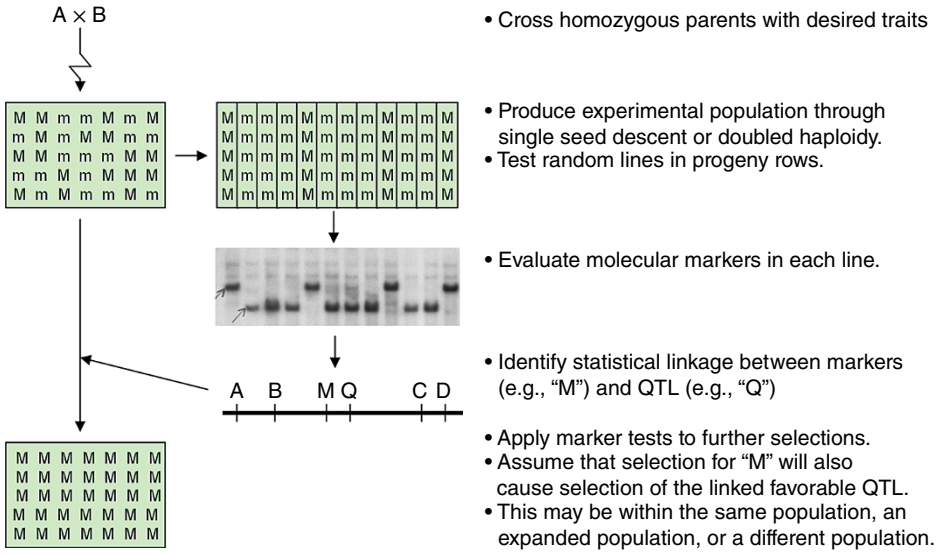


Figure 3.16. A simplified strategy for marker-assisted selection (MAS). Here, a significant association between a QTL (Q) and a molecular marker allele (M) is identified in an experimental population. This information is applied in future populations in order to select Q indirectly through its linkage to M.

can limit the effectiveness (or possibility) of using MAS include (a) some breeding facilities lack the equipment and people trained to apply MAS, (b) incomplete linkage between a marker and a target QTL may reduce the effectiveness of MAS, (c) the marker must be polymorphic on the parents, and (d) MAS is effective only if the alleles being selected are important relative to other alleles in the population. This last factor is the key to the success or failure of every MAS application. It may seem like an obvious statement, but MAS relies on the ability to predict the value of alleles. The quality of those predictions rests on many factors, but a key factor is the behavior of an allele in the presence of other alleles and other physical environments where it has not yet been tested. For example, a breeder might identify that allele A_1 at locus A has a positive effect on yield. But this prediction would be made in a limited set of environments, and with a limited set of germplasm. A breeder who crossed a parent containing allele A_1 with a new parent containing allele A_4, and selected for A_1 using a linked marker, might never discover that allele A_4 is actually better than allele A_1, or perhaps that allele A_1 causes plants to be susceptible to a disease that was not present when A_1 was first characterized. For these reasons, MAS should never be applied independently from phenotypic selection, and most successful applications of MAS have been as an enhancement to phenotypic selection rather than as a replacement.

One technique that is receiving increased usage is called genomic selection. With the increased capacity for identifying and using molecular markers facilitated by high throughput DNA sequencing (see Chapter 7), it is now possible to select across a genome rather than select for one or several markers for single QTLs. In soybean, 50,000 single nucleotide polymorphic (SNP) markers are now available for use by breeders. As well, most of the US soybean germplasm collection has been genotyped using these markers, providing an important reference resource for breeders. Genomic selection is an extension of MAS. Practical use of MAS has been primarily limited to disease resistance genes or QTL with large effects. Genomic selection involves the use of a large number of markers spread relatively uniformly over the entire genome combined with phenotypic data for traits of interest. With an appropriate genotyped and phenotyped reference population, it is possible to generate a genotypic value for each allele at every marker position on the genome. This information



Figure 3.17. Visualization of SNP markers on chromosome-1 for a set of soybean varieties. Each column represents a locus position on the chromosome, and each row represents a different soybean variety. Most loci have two alternate alleles, which are colored to represent the DNA base present in a homozygous state in the corresponding soybean variety. The predicted value of each allele is determined by testing a reference population where phenotypes are known. A predicted genotypic value of each soybean variety is then derived as a summation of predicted allele values, and varieties with the highest overall genotypic values are selected. (*See insert for color representation of the figure.*)

can then be applied to a set of genotyped but otherwise unknown experimental varieties, and to select those with that are predicted to have the best genotypic or breeding values. While the goal of genomic selection and conventional selection is the same, that is, to identify the best line as a new variety, the approaches are different. Conventional selection concentrates on identifying superior lines. Whole genome selection concentrates on identifying high genetic values for loci. The testing program involves training sets of lines where the emphasis is not on the lines themselves but the genetic values for loci that differ among the lines. Genomic selection then integrates genotypic values for the experimental lines and predicts a superior genotype from the pool of experimental genotypes (Fig. 3.17).

3.5.3. Mutation Breeding

Mutations are genetic modifications that occur in the DNA of plants, producing new alleles that are different from the alleles that the plant inherited from its parents. Mutations can be small and localized, or they can cause major structural rearrangements of entire chromosomes. Localized mutations include base substitutions and small insertions or deletions. Because most amino acids are coded by two or more different codons, many base substitutions are "silent," detectable only through DNA sequence analysis. Most mutations that occur in noncoding DNA are also silent, although they can sometimes affect gene expression or chromosome structure. Mutations that cause the transcription of a different amino acid are more likely to cause phenotypic change, most likely through their influence on protein folding or their alteration of an active site in an enzyme.

Fundamentally, the success of all plant breeding depends on mutations that have occurred at some point in the evolution of a species. However, the great majority of random mutations are

deleterious, so breeders rely on a relatively small number of mutations that have been presorted through natural selection because they provide some type of selective advantage in at least one environment. Beneficial mutations that arise naturally are very rare, and may mostly go unnoticed. However, it is possible to artificially induce mutations at frequencies that are much higher than the natural rate. This can be done through radiation (usually applied to seeds prior to planting) or through chemical induction.

The artificial induction of new mutations has been employed when natural variation for a trait is not available. However, it has some disadvantages (see the text); and for these reasons, most breeders consider it to be a technique of last resort.

1. Most importantly, mutations are almost always deleterious, and it is highly unlikely to find a beneficial mutation in a specific gene that affects a specific trait. Therefore, any mutation breeding strategy must be capable of examining large numbers of mutated progeny.
2. Mutations are often not noticeable in a first generation because they are in a heterozygous state. Therefore, breeders must usually look at the offspring of a mutated population.
3. Mutations are usually induced simultaneously in many different genes. Therefore, even if a line is found with a desirable mutation, the same line probably caries many other undesirable mutations that must be bred out.
4. Finally, there is the question of crop safety. It is possible that a new mutation may have unpredictable effects on nontarget traits. This possibility can never be completely ignored; and for this reason, artificially induced mutations are considered by some regulatory systems (including Canada's) to be equivalent to artificial genetic transformation events.

3.5.4. Apomixis

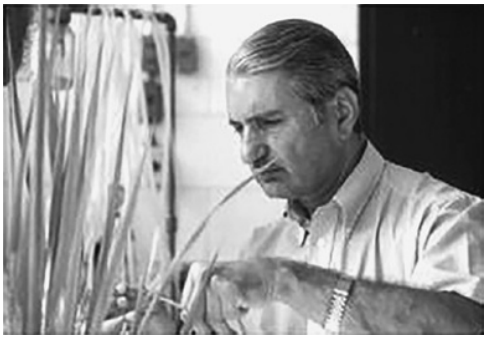
Apomixis is a genetic phenomenon whereby seed is produced without pollination. There are several types of apomixis, but most types result in the production of seed that is identical in genotype to that of the parent plant. Dandelion is a notorious weed that exhibits apomixis; thus all of the seed from a single plant is likely to be identical. Very few cultivated crops exhibit apomixis; these include the turfgrass Kentucky bluegrass, and some lesser-known tropical grasses. However, many research initiatives have attempted, or are attempting, to introduce apomixis into other crops such as maize in order to take advantage of the perpetual hybrid vigor that could be obtained if this were successful. One might speculate that the amount of commercial interest in this endeavor is low, since this would theoretically allow agricultural producers to save their own seed in a crop that might otherwise require the continual purchase of hybrid seed.

3.6. CONCLUSIONS

Plant breeders—who are part scientists, part artisans, part entrepreneurs, part extension workers, and part economists—are a special breed in themselves. Plant breeding is an important field of applied genetics using many of the genetic discoveries made in the previous century, and plant breeders have developed highly scientific approaches to plant variety development. Yet these approaches still leave ample room for personal philosophy, artistic license, and all of the practical challenges of balancing objectives with reality. Over the next century, breeding will incorporate new discoveries and new technologies, but it will almost certainly continue to rely on the principles of sexual recombination and selection.

LIFE BOX 3.1. GURDEV SINGH KHUSH

Gurdev Singh Khush, Former Head of Plant Breeding, Genetics and Biotechnology, International Rice Research Institute, Philippines, and Adjunct Professor, University of California Davis; Winner of the Japan Prize (1987), World Food Prize (1996), and Wolf Prize (2000); Member of National Academy of Sciences and Royal Society of London.



Gurdev Khush. Courtesy of Gurdev Khush.

I was born in a farming family in Punjab, India, in 1935. As I was growing up, I took part in various farming operations and developed an interest in plants. Farm yields were extremely low and poverty was rampant in farming communities. My father was the first person in our village of about 5000 to graduate from high school. He inculcated in me the value of education. I chose to major in plant breeding as an undergraduate at the Government Agricultural College (now Punjab Agricultural University) in Ludhiana and graduated in 1955. Facilities for higher education in India at that time were very limited and I decided to study abroad. I borrowed some money and proceeded to England where I worked in a factory for a year and a half. I returned the borrowed money and saved enough for travel to the United States. I enrolled at the University of California, Davis, in 1957 for a doctorate in plant genetics. I had the good fortune to work under the supervision of a world renowned biologist Professor G. Ledyard Stebbins. After completing my Ph.D. in 1960, I joined the group of another equally outstanding geneticist, Professor Charles M. Rick, as a postdoctoral associate and worked on cytogenetics of tomatoes for 7 years. My solid background in plant genetics proved to be extremely useful in my future career as a plant breeder. In 1966, I was offered the position of a Plant Breeder at the International Rice Research Institute (IRRI), and I moved to the Philippines in August 1967.

The 1950s and 1960s were decades of despair with regard to the world's ability to cope with food–population balance, particularly in the tropics. The cultivated-land frontier was closing in most Asian countries, while population growth rates were accelerating because of rapidly declining mortality rates resulting from modern medicine and health care. IRRI was established to address the problem of stagnant rice yields, the main cause of poverty and hunger in Asia. Conventional rice varieties were tall and lodging susceptible. When nitrogenous fertilizer was applied, those varieties grew even taller, lodged badly, and yields were actually reduced. A breakthrough occurred in doubling the yield potential of rice through reducing the plant stature by the introduction of a dwarfing gene. The first short-statured rice variety IR8 was lodging resistant and highly responsive to nitrogenous fertilizer. It had double the yield potential of conventional varieties. However, it had poor grain quality and was susceptible to diseases and insects. The major focus of my research was to develop improved germplasm with high yield, shorter growth duration, superior grain quality and disease- and insect-resistance. I developed numerous breeding lines with the aforementioned traits. IR36 was the first variety with all the desirable traits. It had high yield potential,

short vegetative growth duration, excellent grain quality, multiple resistance to major diseases and insects, and tolerance to adverse soil conditions such as iron toxicity and zinc deficiency. It was grown on 11 million hectares of rice land during the 1980s. No other variety of rice or any other crop had been as widely planted before. Thirty-four varieties were released under IR designation (IR8–IR74). Seeds of improved breeding lines were shared with national program scientists at their request and through international nurseries. Thus, seed materials were sent to 87 countries irrespective of geographic location or ideology. These materials were evaluated for adaptation to local growing conditions. Some were released as varieties and others were used as parents in local breeding programs. Thus, 328 IR breeding lines have been released as 643 varieties in 75 countries. It is estimated that 60% of the world rice area is now planted to IRRI-bred varieties or their progenies. Large-scale adoption of these varieties has led to major increases in rice production. Average rice yield has doubled from 2 to 4 tons per hectare. Rice production increased from 257 million tons in 1966, to 615 million tons in 2005: an increase of 140%. The price of rice is 40% lower now than in the mid-1960s. This has helped poor rice consumers who spend 50% of their income on food grains. Thus, these IRRI-bred varieties have had a significant impact on food security and poverty alleviation and fostered economic development particularly in Asia, where 90% of the world's rice is grown.

I was fortunate to have had the opportunity to lead one of the largest and most successful plant breeding programs at IRRI. I had a team of motivated plant breeders, plant pathologists, entomologists, and cereal chemists supported by a dedicated Filipino staff. We had a large collection of germplasm, liberal financial support, modern laboratories, and adequate field space. The opportunity to work with scientists in rice-growing countries was another reason for our success. In addition to conventional hybridization and selection procedures, my team employed other breeding approaches such as ideotype breeding, hybrid breeding, wide hybridization, rapid generation advance, molecular MAS, and genetic engineering. I had the opportunity of working with numerous trainees from rice growing countries that came to IRRI for a degree (MSc and PhD) and nondegree training. Upon returning to their countries, they became our valued collaborators. Several of our trainees are now holding positions of leadership in their respective countries. This had a multiplying effect and all the rice-growing countries are now using crop development methodologies and germplasm initially developed at IRRI.

The science of plant breeding is now at a crossroads. Breakthroughs in cellular and molecular biology have added new tools to the breeder's toolbox. MAS has increased the efficiency of selection and reduced the time taken for varietal development. Genetic engineering has permitted the introduction of genes into crop varieties from unrelated sources across incompatibility barriers. In 2006, 102 million hectares were planted to transgenic crops in 22 counties. The science of genomics is likely to improve the efficiency of plant breeding further. The entire genome of rice has been sequenced and efforts are underway to determine the functions of an estimated 40,000 rice genes through functional genomics. Similar efforts are underway in many other crops. Once useful genes for crop improvement are identified, it will be possible to move these genes into elite germplasm through conventional or biotechnological approaches. It is important that plant breeders have a good background in biotechnology and that they work with specialists in the field. The marriage between the ancient profession of plant breeding and the new field of biotechnology will be good for future advances in crop improvement.

LIFE BOX 3.2. P. STEPHEN BAENZIGER

P. Stephen Baenziger, Professor and Nebraska Wheat Growers Presidential Chair, University of Nebraska.



Stephen Baenziger with Dr. Sanjaya Rajaram, World Food Prize Laureate, formerly of ICARDA and CIMMYT looking at an in situ collection of wild barley in Syria, near the origin of barley. Courtesy of Stephen Baenziger.

"Give us this day our daily bread." Although I am not particularly religious, those words have always moved me. When I was in high school, I thought of becoming a human nutritionist so that I could work on world hunger. The US Senate had a subcommittee led by Sen. McGovern on hunger in America that catalogued the dismal state of the poor and Paul Ehrlich published *The Population Bomb* highlighting, quite incorrectly, that massive famines were set to occur in the 1970s. In college, I was a biochemistry major, which was the pre-med major, a group of students whom I never really enjoyed being with because they seemed more interested in their grades than the knowledge (getting into medical school was very competitive), so I gravitated to plant biology, a field that the pre-meds did not know existed.

The professors in plant biology were spectacular (e.g., Winslow Briggs and Lawrence Bogorad) and I decided that, as a nutritionist, I would better define a problem, but not really solve its root causes. Food would still be limiting. Hence, I decided to work on the production side to ensure that there was ample food for those who needed it. At this time, the Green Revolution in wheat, led by Norman (Norm) Borlaug of CIMMYT (a Nobel Laureate for Peace), and in rice, led by Henry (Hank) Beachell, then Gurdev Khush of IRRI (both later became World Food Prize Laureates), had greatly increased the food supply and the predicted famines never occurred. In graduate school, David Glover, who was working on breeding high lysine maize (now referred to as quality protein maize), offered me an assistantship and sealed my fate to become a plant breeder. It was also the last time that I worked on maize.

My first job was to develop small grains (wheat and barley) germplasm with improved disease resistance and tolerance to acid soils (note that I only audited one plant pathology course in graduate school and never took a soils course) for the USDA. Probably, the most interesting aspect of this position, in addition to the excellent scientists that nurtured me, was that the position had been vacant for 4 years and most of the germplasm was transferred or gone. Hence, we needed to rebuild the program from scratch. In winter wheat and barley breeding, it takes 12 years to release a new cultivar and usually at least 8 years to release good germplasm. It was quite clear that time was working against us, so we began a doubled haploid program in hopes we could rapidly inbreed lines and shorten the time to release. Though I have never had sufficient funds to use doubled haploids except for very special genetic studies, this approach is now very common in well-funded commercial breeding programs. Working on germplasm improvement also showed me that despite the massive genetic resources available to wheat and barley, germplasm can be limiting so genetic transformation studies using biotechnology are very important in crop improvement.

After working with the USDA and a short period with Monsanto, I became the small grains (winter wheat, barley, and triticale) breeder at the University of Nebraska. The collaborative USDA-University of Nebraska wheat breeding effort under the stewardship of John Schmidt, Virgil Johnson, Rosalind Morris, and Paul Mattern had been one of the most successful breeding programs in the

United States. At one time 96% of the wheat grown in Nebraska, 40% of the hard winter wheat grown in the United States, and 20% of the wheat grown in the United States came from their program. Here, I learned that breeding can have an impact. I also learned that each crop has special tools that can be used to approach specific scientific questions. While maize had excellent molecular markers, wheat initially had few. However, wheat had chromosome substitution lines (developed at Nebraska by Rosalind Morris) where we could study single chromosome effects across the diverse environments of the Great Plains. In this work, we found that chromosome 3A would increase or decrease grain yield by 15% in the two backgrounds that Rosalind Morris developed. We then used cytological tools to break up these chromosomes by recombination and coupled them with molecular markers to study this chromosome in great detail. In this way, we developed the populations and the phenotypic data while waiting for the molecular marker technology to catch up. It took Rosalind most of her professional career to develop the substitution lines; and after 30 years, we are still studying various aspects of this chromosome because grain yield is still the most important trait in plant breeding. These studies involve huge numbers of lines and the randomized complete block designs were inadequate with the highly variable conditions under which wheat is grown. Working with statisticians, we implemented various statistical methods (nearest neighbor, incomplete block designs) to remove spatial variation in the fields, and to improve our phenotypic estimates. Large experiments require these statistical approaches wherever fields lack uniformity. Genotyping lines has greatly improved and we can and do genotype thousands of lines efficiently at a very reasonable cost. However, nothing can replace good phenotypic data. If a breeder must be knowledgeable in a number of scientific disciplines, and if breeding is built upon the work of previous breeders, perhaps my program has benefited as much or more than most breeding efforts. However, I hope that my ability to ask the pertinent question, to be curious, and to constantly learn/improve how to better measure, combine, and understand the traits that breeders work with has been my contribution. That, my cultivars, and my students will be my legacy.

LIFE BOX 3.3. STEVEN D. TANKSLEY

Steven D. Tanksley, Professor of Plant Breeding & Genetics, Cornell University; President, Nature Source Genetics.



Steven D. Tanksley. Courtesy of Steven D. Tanksley

I grew up in the mountains of Colorado, far removed from the heartland of American agriculture. Despite the short growing season and harsh climate, my mother always managed to coax a garden. It was through her that I came to love plants. Unaware of the many career options in the plant sciences, I headed to college (Colorado State University) with every intention of becoming a doctor. As is the case with most of us, pure chance altered the course of my life. To support myself in college, I took the unenviable job of weekend custodian in the student center. The work was not so bad, but the hours (5 A.M. on both Saturday and Sunday mornings) were terrible. While many of my friends were just coming home from late night revelries, I was pulling myself out of bed to head to work. Seeing the miserable state

of my situation, I was determined to find employment that might afford some level of a social life. I chanced on a job posting for "weekend waterer" in the greenhouse complex managed by the Agronomy Department. The wages were lower than my custodial position, but the hours were much better. It was one weekend when I was watering that I happened to overhear a barley geneticist (Dr. Takumi Tsuchiya) talking in the most animated manner with a graduate student about some topic I could hardly understand. Apparently, they had just discovered something new—something no one had known before. At that moment, it dawned on me that for all my years of study; I had only been reading and memorizing "facts" from books. But, here I was, present at the moment of discovery—where knowledge is created. I was hooked! My watering job gave way to part time lab technician and independent research studies on both barley and sugar beet. Medicine was a fading memory and soon I was headed off to graduate school in genetics at the University of California, Davis.

I arrived at UC Davis with the stars perfectly aligned. Molecular biology was in its infancy and I was in the perfect place to experience and participate in the revolution that would ensue. Studying under the world-renowned geneticist and naturalist, Charles Rick, I was able to apply the new tools of molecular genetics to study natural diversity and build the first "molecular map" (isozymes at the time) of tomato. I also forged a lifelong friendship and collaboration with another graduate student, Dani Zamir. What a blessing that was! Together, we began exploring the implications of these new developments in molecular biology to plant genetics and breeding.

My career path led to faculty positions at New Mexico State University and then Cornell University. Through graduate students, postdocs, and multiple collaborations, we built DNA-marker based maps in a number of crop species—including tomato, potato, and rice. These new maps opened the door to study the genetic basis of quantitative variation. Prior to this time, quantitative genetics was largely a branch of the statistical sciences and not amenable to the same tools by which single gene traits were studied. Using the molecular maps, we were able to identify and map the individual loci (QTLs) responsible for quantitative traits. The molecular maps also opened the door to cloning single gene traits (e.g., disease resistance) and eventually the genes underlying quantitative traits. I was especially interested to determine what specific genetic changes cause quantitative variation. The first QTL we cloned for fruit size in tomato (FW2.2) was quite a surprise. The genetic change was not in the coding part of the gene as expected; but rather, it lay in the regulatory portion of the gene. At the time, many people thought this might be an anomaly—an exception to the rule. But, now with many QTLs having been cloned and studied in a variety of organisms, it is clear that a large portion of natural, quantitative variation is due to regulatory changes, rather than changes in the coding portions of genes.

The "last" stage of my journey has taken me from academia to industry. In 2006, I cofounded a computational genomics company, Nature Source Genetics, dedicated to maximizing the use of natural genetic variation in crop plants by combining the disciplines of mathematics, computer science and genomics. As Director of Research, I am able draw upon my 30 years of experience in academia and fuse it with the amazing power of the computational sciences. As an added bonus, I have now broadened my experience to more than 15 crops species, including field and vegetable crops as well as perennial tree species.

To end on a personal note, I would like to pass on something that my grandfather once told me. He said: "Lucky are the people who don't have to work for a living." I came to understand what he meant is that, if you can find an occupation that captivates your imagination, draws on your strengths, and provides a sense of meaning to your life, then you will never feel like you are "working." In that regard, my life has been a very lucky one indeed!

REFERENCES

Allard RW (1999): *Principles of Plant Breeding*, 2nd ed. John Wiley & Sons Inc., New York.

Barton NH, Keightley PD (2002): Understanding quantitative genetic variation. *Nat Rev Genet* **3**:11–21.

Donald CM (1968): The breeding of crop ideotypes. *Euphytica* **17**:385–403.

Duvick DN, Cassman KG (1999): Post-green revolution trends in yield potential of temperate maize in the North-Central United States. *Crop Sci* **39**:1622–1630.

Fehr WR, Hadley HH (1980): *Hybridization of Crop Plants*. American Society of Agronomy and Crop Science Society of America, Madison, WI.

Frey KJ, Holland JB (1999): Nine cycles of recurrent selection for increased groat oil content in oat. *Crop Sci* **39**:1636–1641.

Hoisington D, Khairallah M, Reeves T, Ribaut J-M, Skovmand B, Taba S, Warburton M (1999): Plant genetic resources: what can they contribute toward increased crop productivity? *Proc Natl Acad Sci USA* **96**:5937–5943.

Orr HA (2005): The genetic theory of adaptation: a brief history. *Nat Rev Genet* **6**:119–127.

Paterson AH (1998): *Molecular Dissection of Complex Traits*. CRC Press, Boca Raton, FL.

Wrike G, Weber WE (1986) *Quantitative Genetics and Selection in Plant Breeding*. Walter de Gruyter, Berlin.

Wright S (1982): The shifting balance theory and macroevolution. *Annu Rev Genet* **16**:1

CHAPTER 4

Plant Development and Physiology

GLENDA E. GILLASPY

Department of Biochemistry, Virginia Tech, Blacksburg, Virginia

4.0. CHAPTER SUMMARY AND OBJECTIVES

4.0.1. Summary

From fertilization to seed to maturity, plants are genetically programmed to grow, develop, and reproduce. Agriculture is greatly dependent on seed production (yield), and yield depends on how plants cope with their environment and other organisms. Since they cannot move around, plants are adept at responding to their environment. They develop and respond primarily by altering their biochemistry, especially in response to plant hormones. In addition, understanding how to manipulate plant development *in vitro* is necessary for the successful engineering of transgenic crop plants.

4.0.2. Discussion Questions

1. Describe the general morphological features of a plant.
2. How is plant fertilization different from animal fertilization?
3. How does the study of mutant plants shed light on gene function?
4. Many genes involved in embryo development also have functions during later stages of development. It has been difficult to clarify these later-stage roles. Why?
5. How do gibberellic acid (GA) and abscisis acid (ABA) physiology affect germinating seeds and mature plants?
6. What is an apical meristem? Name one gene involved in shoot meristem identity, and describe the role it plays during development.
7. What is etiolation?
8. How do the PHY proteins function as light receptors?
9. How do the quiescent zone and root cap structures and properties differ?
10. How do guard cells participate in photosynthesis and respiration?
11. How do the adaxial and abaxial surfaces of the leaf differ?
12. You have isolated a gene whose expression is confined solely to the developing leaf primordia and have obtained a loss-of-function mutant for this gene. Speculate as to what phenotype might result in this mutant and also explain the basis for your speculation.

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

13. What is the difference between a daylength-neutral and a long-day plant?
14. Describe the ABC model of flower development and speculate as to what phenotype would result if a C function gene were overexpressed in all whorls.
15. In snapdragon, Floricula mutants contain shoots with the characteristics of an inflorescence meristem in place of a floral meristem. Hypothesize what the wild-type function of the Floricula gene is, and speculate as to why investigators are interested in overexpressing this gene in aspen trees.
16. Describe the major effects of plant hormones on growth and development. Also describe the connection of the 26S proteasomal pathway to signaling via various plant hormones.

4.1. PLANT ANATOMY AND MORPHOLOGY

Before considering the developmental and physiological processes that can impact plant biotechnology, one should have some basic knowledge of plant anatomy and morphology. This section is designed to provide a closer look at internal structures and cells within the plant.

Most plants are composed of the shoots (or aboveground tissues) and roots (or the belowground tissues). See Figure 4.1. The shoot apex consists of the topmost tissues of a seedling or plant and contains the shoot apical meristem (SAM) and the developing leaves or leaf primordia. The SAM is a dome-shaped region of dividing cells at the tip of the stem (Fig. 4.1). The SAM is the control center of the plant and directs the development of all aboveground differentiated tissues such as the stems, leaves, thorns, flowers, and fruits. Cells within meristems undergo cell division quickly, and these are usually smaller because they have smaller vacuoles than differentiated plant cells (Fig. 4.1).

The root also contains a similar control center, the root apical meristem (RAM) that functions in generating new root cells within the root tip (Fig. 4.1). A section through the root shows that roots are often full of starch granules that can be visualized by staining with potassium iodide, which turns starch a blue-brown color. One can also see the meristematic zone at the root tip, the root cap, a

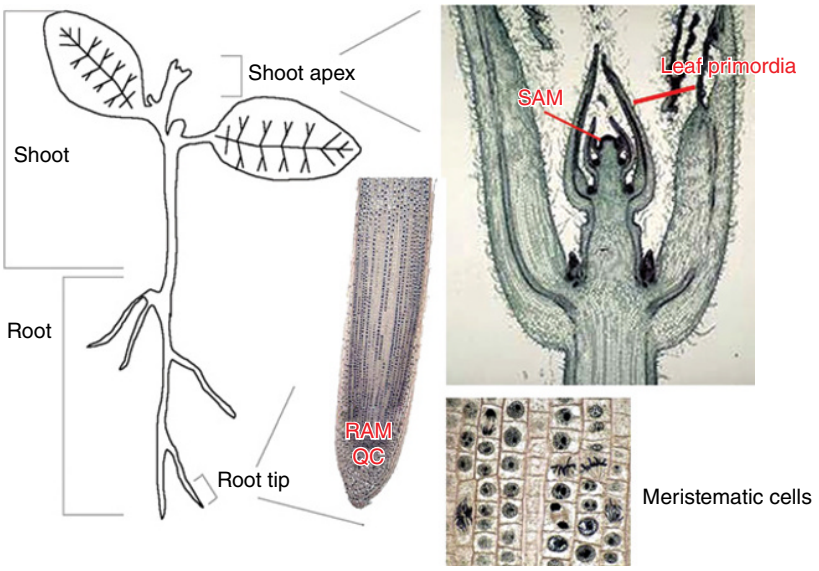


Figure 4.1. Plant anatomy and morphology. The seedling shoot and root systems are indicated, as are the shoot apical meristems (SAM) and root apical meristems (RAM), tissues that direct the major growth and differentiation of plants. Active cell division within the meristem is shown in the last panel; note the presence of two nuclei in some cells. (Adapted with permission from Dr. Dale Bentham.)

protective covering, and the ordered files of cells resulting from the root initial cells within the RAM. One may also be able to view the quiescent center (QC), so called because cells are "sleeping" or slow to undergo cell division.

Axillary buds are the third type of meristems that give rise to new tissues. Axillary buds may be found on stems and, under the right conditions, can give rise to new SAMs.

Plant cells within shoots and roots are organized into specialized tissues that enable the organism to carry out necessary functions. The tissue systems of plants are the dermal, vascular, and ground tissue systems. The dermal system is composed of the epidermal, or outermost, cell layer, which covers the entire plant. The vascular tissue system is composed of the xylem, phloem, and other conducting cells that transport water and nutrients. This tissue is present in most plant tissues, but it can be arranged differently within each organ. The ground tissue is composed of the cells in between the epidermis and the vascular tissue.

There are many different specialized plant organs. In addition to the SAMs and RAMs, most angiosperms contain stems, leaves, lateral roots, and reproductive tissues such as flowers and their component tissues (anthers, filaments, pollen, etc.). Each of these tissues can impact the development and physiology of the plant, and as such must be considered when manipulating gene expression in transgenic plants. Specific considerations for each of these tissues will be discussed as we chart the development and physiology of an average plant in the succeeding sections.

4.2. EMBRYOGENESIS AND SEED GERMINATION

4.2.1. Gametogenesis

The lifecycle of flowering plants alternates between a haploid organism, the gametophyte, and a diploid organism: the sporophyte. Plants have male and female gametophytes, both of which are multicellular and are produced within the flower (Fig. 4.2). The mature male gametophyte, the pollen grain, has three cells: a vegetative cell and two 1*N* sperm cells. Pollen development (Fig. 4.3) occurs in the anther, which is a specialized structure of the flower, with the meiotic divisions of the microsporocytes to form a tetrad of haploid spores. The microspores are embedded in callose, and

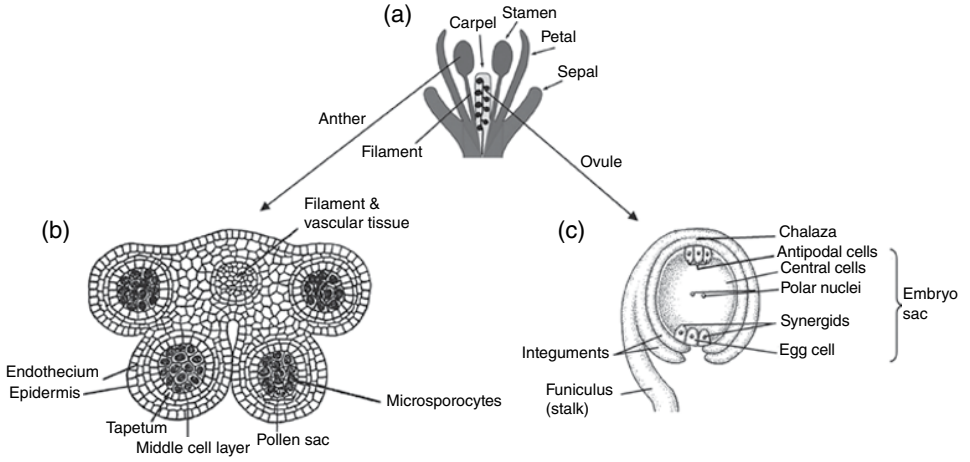


Figure 4.2. Gametogenesis. Schematics of (a) an *Arabidopsis* flower with the floral organs identified, (b) a cross section through the male organs (anther, filament) showing the site of male gamete formation, and (c) the female ovule contained within the carpels of the flower showing the site of female gamete development. (From Wilson and Yang (2004). Reproduced with permission from Society for Reproduction and Fertility.)

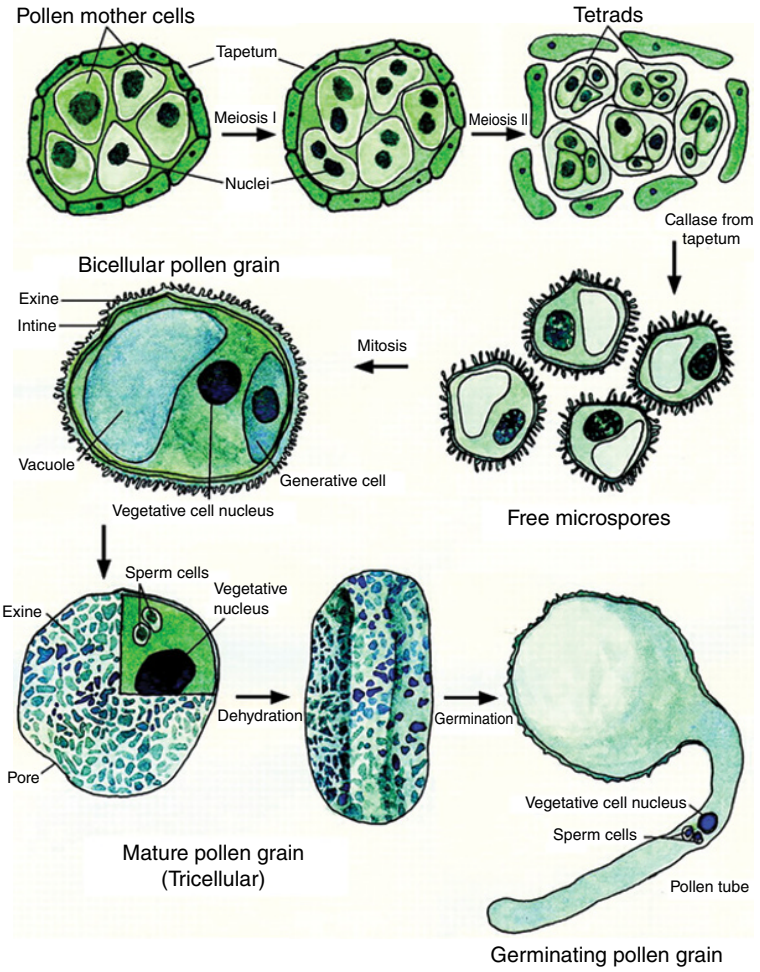


Figure 4.3. Pollen development. (From McCormick (2004). Reproduced with permission from the American Society of Plant Biologists.)

release from the tetrad requires enzymes secreted by somatic cells in the anther. Mature pollen grains have complex walls with two layers: the inner intine and the outer exine layer.

Self-incompatibility as a mechanism to limit reproduction was discussed in Chapter 2. However, fertilization also depends on the gene products that are required for normal development of the *pollen* and *ovules*. Scientists have identified several of these required gene products by taking a genetic approach. To identify molecules involved in either *gameteogenesis* (the formation of gametes) or *fertilization*, geneticists have utilized mutant populations of the model plant called *Arabidopsis thaliana*. The genome of *Arabidopsis* is fully sequenced, and many different mutant populations containing a loss of function in individual genes are available. The first mutant collections were most often composed of plants containing random, single base pair mutations, or T-DNA insertions. Both types of mutant collections can be screened, and mutants identified on the basis of the phenotype. For example, mutants defective in a gene required to form a female or male gamete will give rise to mature plants with low fertility. Low fertility can be somewhat easily scored in a random mutant population by looking for low seed set. In *Arabidopsis*, the seeds are produced within small, elongated fruits called *siliques*. Finding a plant with fewer siliques, or empty siliques, is an indication that a mutant has lost the function in a gene required for gamete development. One can determine which gamete has been affected by examining the appearance of both female and male gametes from the

candidate mutant plant. For example, if pollen grains appear normal and germinate pollen tubes *in vitro*, then most likely, the defect is *not* in male gametophyte development. The scientist would then examine the appearance of the female gametophyte within the flower. Outcrosses of the candidate mutant pollen to a wild-type pistil, and the reverse outcross (candidate mutant female to wild-type male cross), can also be important in determining which gamete is defective (Wilson and Yang 2004; Boavida et al. 2005).

Using such screens and outcrosses, geneticists have isolated several genes required for pollen and ovule development. In *pop* mutants, for example, the *exine layer* of the pollen grain does not develop properly, resulting in altered hydration of pollen grains. Without normal hydration, the pollen tube guidance is not normal, and fertilization is greatly lowered. These mutants point to the idea that structural components of the pollen grain itself are important for male fertility.

Female sterile mutants have also led to the identification of genes required in female gametophytic development. ANT, BEL1, SIN1, and ATS gene products were each identified in mutant screens. Each of these genes encode proteins required for ovule development. For example, the *ant* mutant cannot make the *integuments* that surround the developing egg cell. ANT is a transcription factor that controls organ initiation and promotes cellular divisions during development of the integuments. ANT is known to interact synergistically with SEUSS (SEU), a transcriptional co-regulator of organ size in flowers. The *bel1* mutant is also defective in integuments, but it does develop a collar of tissue that surrounds the egg cell. Thus *bel1* mutants have an altered integument and the function of the BEL1 protein is to specify integument identity within the developing female gametophyte. BEL1 is known to be a transcription factor and is thought to control cytokinin and auxin signaling during ovule development. The *sin1* mutant also has altered integuments that are shorter. This mutant is of special interest in that the SIN1 protein is a homolog of the DICER protein that functions in generating small, interfering RNA molecules (siRNA) that suppress *gene expression* at the posttranscriptional level. The fact that a DICER-like enzyme is required for normal ovule development strongly suggests that posttranscriptional regulation of ovule identity genes is important for maternal development.

The phenotypes of these mutants help build a model of the ovule developmental pathway. They suggest that during the process of flowering, the ovule primordia initiate and then gain ovule identity. For example, primordia initiation must include ANT function, which is then followed by the action of genes that specify the integuments like BEL1. In this model, SIN1 function would follow, giving rise to the normal shape and size of the integuments. Thus, by using a combination of genetic and molecular approaches, developmental biologists can order gene function in the development of specific tissues.

4.2.2. Fertilization

The beginning of a plant's life starts with fertilization of the haploid (1*N*) egg cell within the ovule by one of the two haploid sperm nuclei carried by the pollen tube of the pollen grain (see Chapter 2 and Fig. 4.2). Development will produce a 2*N* plant embryo surrounded by maternal tissues within the carpels. Plants actually undergo a separate fertilization event that creates the 3*N* *endosperm*. The endosperm results from fusion of the other 1*N* *sperm* nuclei with the two *polar nuclei* (2*N*) within the central cell of the ovule. The resulting endosperm tissue can transfer nutrients into the developing embryo. Thus plants, like animals, have a food supply handy for the developing embryo. The triploid nature of the endosperm has been speculated to be a mechanism for controlling gene dosage or a way for maternal control of embryo development (Berger et al. 2006). An interesting phenomena called *endoreplication*, or *endoduplication*, occurs at an increased rate within the endosperm. This process involves DNA replication in the absence of cell division, resulting in a high *N* number within certain cells of the endosperm.

Studies on *Ephedra trifurca*, a nonflowering seed plant that is a close relative of the angiosperms, have revealed key differences in fertilization. This plant, from which Mormon tea is made, has a

second fertilization event that leads to formation of a second embryo instead of endosperm development. This difference has prompted speculation that the modern endosperm of today's plants may have evolved from a second embryo like that found in *Ephedra*. We know that fertilization and development of the embryo and endosperm in angiosperms are dependent on each other; that is to say that normally the endosperm *must* develop in order for the embryo to develop. However, there is a mutant that has been identified where fertilization of the endosperm occurs in the absence of embryo fertilization and development. This mutant, called *fie* (*fertilization-independent endosperm*), suggests a connection between endosperm development and chromatin as the FIE gene product is a type of *polycomb* protein. Polycomb proteins were first discovered in the fruitfly, *Drosophila melanogaster*, and act by "locking" chromatin into accessible or nonaccessible forms that dramatically alter gene expression in the next generation. Thus, the FIE polycomb gene product may be necessary to "lock in" the appropriate chromatin pattern for the communication between the embryo and the endosperm developmental processes (Twell 2006).

4.2.3. Fruit Development

Fertilization is also important to consider in plant biotechnology as it directly impacts the process of fruit development. Fertilization is the trigger for growth of the ovary that can then develop into a fruit. The term *fruit* can be used to describe any ovary that initiates a growth program after fertilization. For example, the enlarged ovary under a decaying rose flower is called a *rose hip* and, like citrus fruits, contains high levels of vitamin C. Fruit development is a strategy thought to attract animals that will eat the fruit and disperse the seeds far from the plant. Animals and plants have coevolved, with animals trying to get the most nutrients (through digestion) from the fruit and seeds, and the plant evolving processes designed to facilitate seed dispersal in contrast to seed digestion. This coevolution may account for the incredible diversity of fruit and seed types.

Fruit development requires both fertilization and growth of the embryo within the seed; thus seed and fruit development are related. For example, in some species lopsided fruit will result when fertilization of ovules on one side of the ovary is defective. The seeds developing from fertilized ovules are thought to signal to the surrounding fruit via their production of growth hormones, such as auxin and cytokinin. There are physiological conditions, however, that will override the requirement for these seed-derived hormones. The process of fruit development in the absence of seed development is called *parthenocarpy*, which is a desirable trait for certain fresh fruit. Some commercial "seedless" varieties, like the seedless watermelon, actually have very tiny, partially developed seeds. In contrast, certain true seedless grape varieties undergo parthenocarpic fruit development in the absence of fertilization of the ovules. Studies on parthenocarpic fruit will lead to a better understanding of the processes that accompany fertilization. One useful tool is the *fwf* (*fruit without fertilization*) mutant from *Arabidopsis*, which is a facultative parthenocarp, setting seed in a normal way when pollinated, but also forming short seedless fruit when left unpollinated. The *fwf* mutant has a mutation in Auxin-Response Factor8 (ARF8), resulting in the uncoupling of fruit development from pollination and fertilization. ARF8 normally acts as an inhibitor to stop further carpel development in the absence of fertilization and the generation of signals required to initiate fruit and seed development.

4.2.4. Embryogenesis

As described earlier, *embryogenesis* begins after the 1*N* egg cell and 1*N* sperm nuclei fuse together, forming a 2*N* embryo. Plant embryogenesis differs significantly from animal embryo development in its lack of cell migration and substantial cell specification. For example, the mature plant embryo within the seed does not contain cells specified to become flower cells or gamete-producing cells. These differentiation events will occur later in development, well after seed germination. Instead, plant embryogenesis will result in the acquisition of bilateral symmetry, an apical/basal or shoot/root axis, and the three types of tissue.

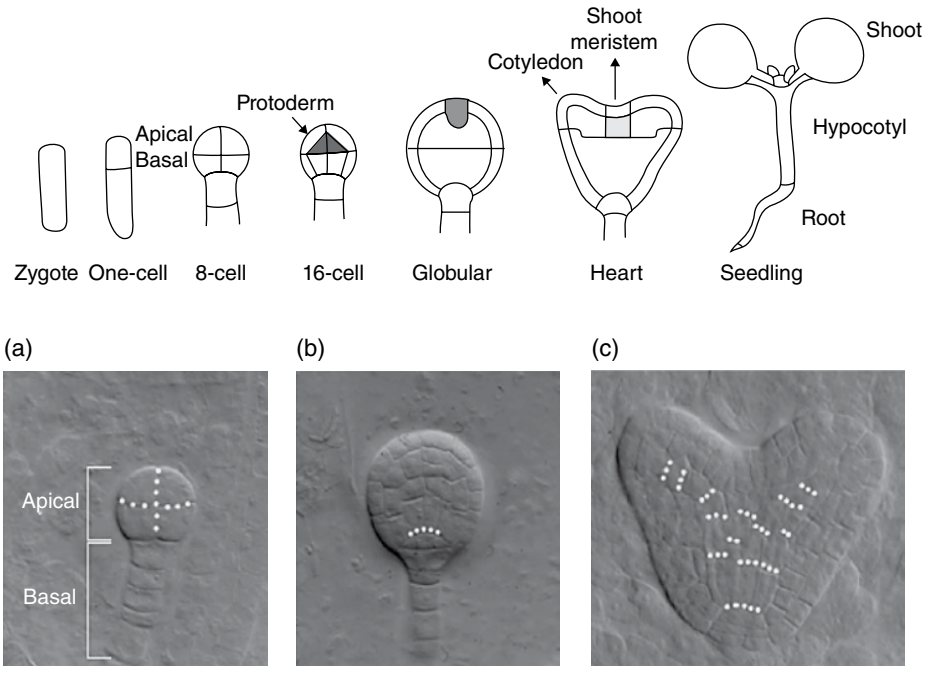


Figure 4.4. Embryo development. (a) Schematic of embryo stages. (b) Scanning electron micrograph of *Arabidopsis* embryos in the globular and heart stages. The white lines indicate the cell division planes. (From Costa and Dolan (2000). Reproduced with permission from Elsevier Science Ltd.)

The first cell division of the plant embryo results in an asymmetric division giving rise to a small upper, terminal cell and a larger, lower basal cell (Fig. 4.4). This establishes a longitudinal or an apical/basal axis in the embryo (Weijers and Jurgens 2005). The upper cell always gives rise to the embryo proper, while the lower cell gives rise to the *suspensor* and the *hypophysis*, forming part of root meristem, root initial cells, and root cap. The suspensor is a highly specialized and terminally differentiated tissue that connects the embryo to the embryo sac and maternal ovule tissue. It functions as a conduit for nutrients and senesces after the heart stage of embryo development. This short-lived unique organ consists of only 7–10 cells in total in *A. thaliana*.

The upper cell of the two-cell embryo undergoes two more cell divisions, passing through the 4- and 8-cell stages, in which a gain of embryo mass occurs. Further cell divisions result in mass of cells on top of the suspensor referred to as the *globular stage embryo* (Fig. 4.4). More cell divisions result in development of the heart stage embryo, so called because of the characteristic heart shape of the embryo. This heart shape results because differentiation of cells has occurred, with some cells beginning to acquire SAM identity in the cleft of the heart, and two lateral domains giving rise to cells destined to form the cotyledons of the embryo. In addition, the RAM becomes specified at this stage. With the development of the SAM, RAM, and cotyledons, the embryo is now beginning a change to bilateral symmetry.

After the heart stage, organ expansion and further cell divisions result in the lengthening of the embryonic cotyledons into the "torpedo" stage (Fig. 4.4). At this point, two patterns have been established: the apical/basal patterns, which allows for shoot versus root development; and the radial pattern, which gives rise to the three types of tissue: (a) the *protoderm* (which gives rise to the epidermis), which divides anticlinally; (b) a middle layer, the *ground meristem* (which gives rise to the cortex and endodermis); and (c) an inner layer, the *procambium layer* (which gives rise to the vascular tissue) (Willemsen and Scheres 2004).

The last stage before the mature embryo stage is the "walking stick" stage, so called because the developing cotyledons have folded down over the SAM. To mature, the embryo must enter a dehydration phase in which metabolism pauses. In the dehydrated state, the embryo within its seed coat is waiting for the appropriate environmental conditions suitable for seed *germination*. The plant hormone *abscisic acid* (ABA) is required for initiating dehydration and establishing seed dormancy. Without an ABA source or a functioning ABA signal transduction pathway, embryos can germinate "precociously" inside a fruit. Thus the study of ABA signaling pathways and the genes turned on by these pathways is directly relevant to the understanding and manipulation of seed germination.

4.2.5. Seed Germination

Germination is the process wherein the embryo imbibes water and returns to growth after dormancy. *Imbibition* is the uptake of water by the embryo within the seed. During this process, the embryonic tissues are loosened and the seed coat usually splits, allowing more water to penetrate the embryo. Once the embryonic cells are rehydrated, the metabolic processes of germination can begin.

Several common requirements are shared by very diverse types of seeds, including temperature and moisture. Some seeds have a light requirement, and some also require a cold pretreatment called *stratification*. These processes promote the increase and/or action of a plant hormone called *gibberellic acid* (GA). GA action is generally considered as antagonistic to ABA, and it is also considered to be the dormancy-breaking hormone. One well-characterized action of GA is the induction of α-amylase production that breaks down stored starches in grain seeds. Germination can occur underground (in the dark) or aboveground (in the light). Either way, the major result of germination is the expansion of the already preformed embryo (Koornneef et al. 2002).

4.2.6. Photomorphogenesis

Imbibition of a seed allows dormant cells to expand and for new cell division to occur within the embryo. The specific type of growth is influenced heavily by the presence or absence of light. Light is the most influential signal from the environment that plants perceive. When a seed germinates aboveground, or in the presence of light, it immediately responds to light with an elegant and complex developmental response called *photomorphogenesis*. If a seed germinates underground or in the absence of light, it undergoes a brief and specific developmental pathway called *skotomorphogenesis*. The purpose of this dark developmental pathway is assumed to be the alteration of growth in the seedling that increases its chance of encountering light, a signal required for the further development of the seedling.

When germination occurs in the dark, the seedling develops into what is called an *etiolated* seedling, which is characterized by increased hypocotyl growth, an apical hook (in dicots), unexpanded cotyledons, and no chlorophyll synthesis. These adaptations to dark can allow for the elongating hypocotyl to push the SAM and cotyledons up through the soil to encounter light. The apical hook thus can protect the new SAM, and chlorophyll synthesis is not needed until light is encountered.

When the seedling encounters light, the elongation of the hypocotyl slows, the apical hook uncurls, and the *cotyledons* expand and begin to assemble functional chloroplasts containing chlorophyll. Transcription of genes encoding the chlorophyll a-/b-binding proteins and part of the Rubisco complex are rapidly upregulated. Thus, if a seed germinates in the presence of light, its *hypocotyl* will be much shorter than that of an etiolated seedling. The apical meristem will then give rise to the first pair of true leaves that differ in structure from the cotyledons and contain *trichomes*, or hairs.

The light receptor required for red light signal transduction is called *phytochrome*, which is composed of an open-chain tetrapyrole pigment called *phytochromobilin* and a protein dimer of 240 kDa. This pigment/protein complex allows for the perception of red light by absorption of either red or far-red light. Phytochrome is distributed throughout many different cell types in the plant, and

more recent evidence suggests that it traffics from the cytosol to the nucleus in response to light, where it interacts with transcription factors such as PIF3 to influence gene expression. Many of the gene products required to construct an active photosynthesizing chloroplast are controlled by the presence of light, and thus are most likely under the control of phytochrome-mediated signal transduction pathways. Phytochrome itself is encoded by five different *Phy* genes called *PhyA* through *PhyE*.

Mutants defective in photomorphogenesis have been instrumental in identifying genes required for this process. There are two general categories of photomorphogenesis mutants: (a) *hy* and (b) *cop* and *det*. The *hy* (*hypocotyl elongated*) mutants look partially etiolated even when grown in the light, indicating that the HY gene products function in the perception of light. These screens identified some of the *Phy* genes and other positive regulators of photomorphogenesis such as HY5, a key transcription factor. In contrast, the *cop* (*constitutive photomorphogenesis*) and *det* (*deetiolated*) mutants were identified by virtue of their light-grown phenotypes when grown in the dark. Many of the *cop* mutants encode proteins that form a large complex called the *COP9 signalsome* (CNS), a nuclear complex that is similar to the 26S proteasome proteolytic complex that degrades ubiquitinated proteins (Rockwell et al. 2006). Both COP and DET act as repressors of light signaling, via regulation of the ubiquitin–proteasome system. COP1 is a RING E3 ubiquitin ligase that targets key regulators for degradation, and DET1 complexes with COP10 and an adaptor protein that is proposed to aid in COP1-mediated degradation.

The lack of etiolation in some *cop* and *det* mutants can be reversed by adding the plant steroid hormone *brassinolide* (Br) (Zhu et al. 2013), suggesting a role for Br signal transduction in photomorphogenesis. The *det2* mutant of *Arabidopsis* has sequence homology with mammalian steroid 5α-reductases. This suggests that the DET2 gene product participates in Br synthesis. Thus, light may control photomorphogenesis by downregulating Br production.

Blue light is another important stimulus for photomorphogenesis and for phototropism (growth toward light). Blue light is perceived by two types of flavin-containing proteins: *crytochrome* and *phototropin*. Both cryptochrome (CRY) and phototropin are encoded by two genes in *Arabidopsis*. CRY proteins appear to function in the nucleus, although there are indications that there may be some CRY functions in the cytoplasm as well. Evidence suggests that phytochrome and cryptochrome physically interact. CRY protein can be phosphorylated *in vitro* by the protein kinase activity of PHY. In addition, PHYB and CRY2 interact in plant extracts. CRY1 and CRY2 also appear to directly interact with COP1, the negative regulator of photomorphogenesis in the dark.

4.3. MERISTEMS

Plant meristems are dynamic structures whose functions are to renew themselves and to give rise to new cells with a different identity. There are three types of meristems: apical meristems, including SAMs and RAMs; the lateral meristems, including the vascular and *cambial* meristems responsible for secondary growth; and the *intercalary meristems*, common to the grasses that occur at the bases of nodes. The common function of these meristems is regulation of cell division that creates new cells specified to become different cell types and renewal of the meristem itself.

4.3.1. Shoot Apical Meristem

Apical meristems are extremely important in terms of growth regulation of plants. As alluded to previously, the SAM gives rise to the aerial parts of higher plants by continuously initiating new organs. The basis of this activity is its ability to maintain a pool of pluripotent stem cells, which are the ultimate source of all tissues of the shoot. The SAM typically consists of a dome of cells connected to two developing leaf primordia (Fig. 4.1). This area contains around 100 cells in *Arabidopsis*. The dome structure contains the least differentiated cells and consists of three different histocytological zones (Fig. 4.5). The central zone in the middle of the dome contains cells that

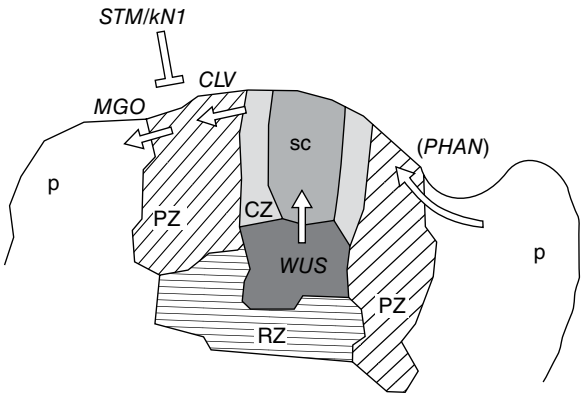


Figure 4.5. The shoot apical meristem (SAM). Schematic of a SAM showing the central (CZ), peripheral (PZ), and rib meristem (RZ) zones. Proteins involved in meristem development are also shown. (From Lenhard and Laux (1999). Reproduced with permission from Elsevier Science Ltd.)

divide infrequently, yet this is the location of the self-renewing undifferentiated stem cells. Surrounding the central zone is the peripheral zone, where the rate of cell division is higher and cells contribute to the organs of the plant, including leaves, inflorescence meristems, and floral meristems. Below the central zone is another region of rapidly dividing cells, called the "rib" meristem. Division and elongation of rib meristem cells gives rise to the stem of the plant.

The SAM also consists of different cell layers. The surface layer of cells is called the *L1 cell layer*. Cells in L1 divide only by forming anticlinal cell walls, that is, cell division is always perpendicular to the meristem surface. As a result, cells in the L1 layer and their daughter cells always remain in this layer. The *L2 cell layer*, below the L1 cells, divide the same way. The *L3* or *corpus cells*, divide in all planes, and fill the interior of the SAM.

A major issue in plant biology concerns how shoot meristems are organized and how molecular information in the SAM determines the precise placement/function of cells. More recent molecular studies indicate that the maintenance of stem cell function depends on a feedback loop involving the CLV1–3 (*Clavata*) gene products and WUS (Wuschel). In *clavata* mutants, the meristem is enlarged, due to excessive accumulation of stem cells, suggesting that CLV1–3 are required to regulate the number of stem cells in the meristem. In contrast, *wus* mutants contains a smaller meristem with differentiated cells, suggesting that WUS is a positive regulator of stem cell identity. Analysis of the interactions between these key regulators indicates that (a) the *Clv* genes repress WUS at the transcript level and (b) WUS expression is sufficient to induce meristem cell identity and the expression of the stem cell marker CLV3. As the different *CLV* genes encode a receptor and a ligand that binds this receptor, it appears that the CLV gene products together form a signal transduction pathway that limits the expression region of WUS. It is known that CLV3 is a secreted peptide signal that binds to CLV1 or CLV2, and this signaling regulates WUS transcription to affect the balance of stem cell differentiation and proliferation in the SAM. Thus the interaction between CLV and WUS maintains stem cell function and the maintenance of the meristem as a source of cells for the shoot.

Other SAM regulatory genes are known to be expressed in the SAM. The *shoot meristemless* (*Stm*) homeodomain transcription factor gene is required for normal SAM function, as *Arabidopsis stm* mutants lack a functional meristem. Further, transgenic tobacco plants expressing an extra copy of the corn *KN1* (STM-related) gene develop superficial SAMs on leaves, suggesting strongly that *KN1* expression directs SAM formation. The *Mgo* genes also play a role in SAM function. The *mgo1* and *mgo2* mutants contain disorganized SAMs and fewer leaves 10 days after germination, suggesting that the SAMs of the mutants delegate fewer cells to the leaf primordia. The *Mgo* genes encode proteins similar to asymmetric cell division regulators in animal cells, suggesting a key role for the MGO proteins in meristematic cell divisions. Finally, the *Phantastica* (*Phan*) genes

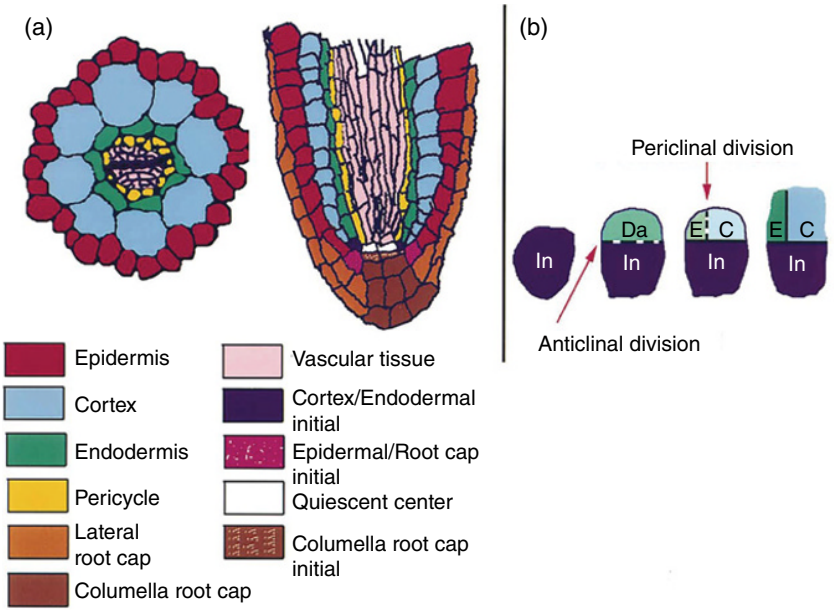


Figure 4.6. Root development. Arrangement (a) and division plane (b) of cell types within the developing root. (From Di Laurenzio et al. (1996). Reproduced with permission from Cell Press.) (*See insert for color representation of the figure.*)

help specify adaxial leaf identity, and thus are involved in leaf primoridia differentiation (Traas and Bohn-Courseau 2005; Shani et al. 2006).

4.3.2. Root Apical Meristem and Root Development

Root development is illustrated in Figure 4.6. Organization of the RAM involves fewer cells than does development of the SAM. The basic organization of the SAM and RAM are similar in terms of having a central region of slowly dividing cells surrounded by cells with a higher cell division rate. Recall that specification of the RAM occurs during the embryonic heart stage. Thus at the heart stage, the radial organization of tissues is in place and the RAM initials and central cells that will generate and maintain the root in the seedling are specified. The QC is the region of slowly dividing cells within the RAM. The QC is involved in RAM activity and maintenance. In bindweed (*Convolvulus arvense*) the QC cell cycle is 430 h, whereas in other cells it is about 13 h. Therefore, the QC must be viewed as the ultimate source for new cells, but not the factory that produces them.

Below the QC lies the *columnella root cap initial* cells, which give rise to the root cap, a protective structure. Above the QC are the epidermal initials, which will form the epidermis and lateral root cap. The cortical and endodermal initial cells give rise to the cortex and endodermis; the final layer is the vascular tissue. The initial cell that gives rise to either endoderm or cortex divides anticlinally once and then periclinally once before these identities are laid down. The portion of the root enclosed by the endodermis is often referred to as the *stele*.

Cell divisions from these initial cells follow a strict pattern of progressive differentiation, resulting in an expansion (elongation zone) and a differentiation (maturation zone) to build a regular arrangement of cell files within the root body. It is not surprising that expression domains of regulatory genes are responsible for cell fate patterning in the RAM. For example, the stem cell population in the RAM is maintained by confining the expression of Wuschel-Related Homeobox5 (WOX5), a homeobox transcription factor, to the QC. Repressor of Wus1 (ROW1) acts in the SAM, but was recently found to limit WOX5 expression outside the QC, by modifying chromatin within the WOX5 promoter (Zhang et al. 2015).

Other important root identity genes are *short root* (*Shr*) and *scarecrow* (*Scr*), which help specify the endodermis and cortical identities of cells, respectively. SHR and SCR proteins function in a novel signaling pathway to determine radial patterning in the root. The SHR protein is translated in the stele and then moves to the adjacent cell layer, where it activates SCR transcription and initiates endodermal specification. The SCR protein is then thought to regulate the asymmetric cell division that results in the formation of cortex and endodermis.

The plant hormone *auxin*, or indole acetic acid (IAA), is required for the formation of the embryonic root, lateral roots, and maintenance of the cellular organization around the initials of the seedling root. Auxin moves through the plant from the shoot, where it is synthesized, to the root using a system of influx and efflux carriers localized asymmetrically in the cells of the vascular tissues. It has been shown that the family of auxin transporters encoded by the *Pin* genes are the auxin efflux carriers and that PIN1 localization becomes progressively polarized in developing embryos. By the globular stage, PIN expression is confined to the basal portion of the embryo, and as embryogenesis proceeds, PIN becomes further localized to the developing vasculature. The effects of auxin on root patterning can be visualized in transgenic plants containing five copies of an auxin-responsive gene promoter element to drive expression of the GUS (β-glucoronidase) reporter gene. When expressed in transgenic *Arabidopsis*, one can visualize auxin content by utilizing an assay that detects GUS activity. The results show that there exists an expression maxima in the root initial cells, supporting the role of auxin in root patterning. Root meristems are the focus of much research (Costa and Dolan 2000; Campilho et al. 2006).

The formation of lateral and adventitious roots also requires auxin. Lateral or secondary roots originate from the pericyle, a specific cell type contained within the stele of the root. Cells within the pericycle undergo cell division, and then further cell division and cell expansion results in the formation of a lateral root. These cells begin cell division in response to auxin and environmental cues and must establish a connection to the vascular trace of the primary root. Adventitious roots can develop from the stems of some plants when placed under inducing conditions. Tomato, for example, can develop many adventitious roots from a cut stem when placed under humid conditions.

Root hairs are another type of cell contributing to the overall root function of absorption of water and minerals. The outer, epidermal layer of the root gives rise to root hairs. Root hair formation occurs within a specific region of the root, a short distance above the region of root elongation. Root hairs are short and short-lived and develop on both primary and secondary roots. Interestingly, a root hair is a single cell that consists of a thin cell wall, a thin lining of cytoplasm that contains the nucleus, and a large vacuole-containing cell sap.

4.4. LEAF DEVELOPMENT

4.4.1. Leaf Structure

Leaves are specialized structures responsible for most of the photosynthesis that takes place in the plant, as well as functioning in respiration and transpiration. Leaves are initiated as primordia from the SAM as described earlier. As leaf primordia are specified by gene such as the *Phan* gene, the abaxial (or top), and adaxial (bottom) surfaces develop (Fig. 4.7a). Recall that leaves differ from the cotyledons in several ways including the presence of the single-celled trichomes, or leaf hairs that function in the secretion of various compounds that can attract or repel insects (Fig. 4.7b).

A cross section of a mature leaf shows the main cell types in the leaf (Fig. 4.7c). The outer epidermal cell layers are derived from the L1 layer of the SAM in both monocots and dicots and do not contain chloroplasts. The exception to this is the *stomatal pore*, which is created from two guard cells that contain a specific number of chloroplasts, depending on the ploidy of the plant. The interior

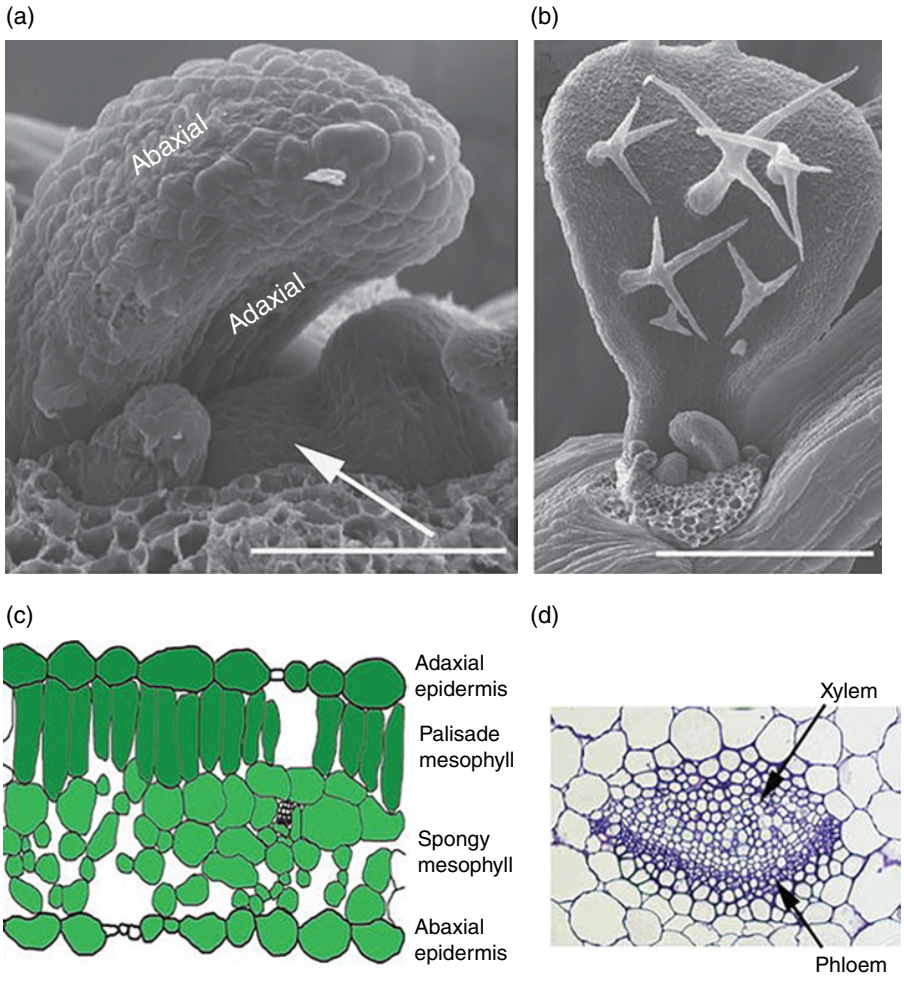


Figure 4.7. Leaf development. (a, b) Scanning electron micrographs of leaf primordia. Note the presence of trichomes in (b). (c) Schematic of leaf cross section showing the different leaf cell types. (d) Cross section through leaf vascular tissue. (From Byrne (2006). Reproduced with permission from the Public Library of Science.)

leaf cells are filled with chloroplasts that will autofluoresce when viewed under a fluorescent microscope. Dicot leaves have a distinct dorsiventrality with an upper (adaxial) layer of oblong palisade cells, and a lower (abaxial) layer of spongy mesophyll cells (both are derived from the L2 layer of the SAM). Vascular bundles containing xylem and phloem are present in the middle or L3-derived layer. Monocot leaves vary, but all contain a single photosynthetic cell type, the mesophyll, and a specialized bundle sheath surrounding the vascular tissue. There are many specializations of leaves, such as in xerophytic leaves, which are adapted to dry conditions, which contain different cell types and arrangements of these cell types (Byrne 2006).

Mature leaves are often surrounded by a waxy, cuticle layer that provides protection and prevents water loss. The epidermal cells secrete this layer and themselves provide protection to the internal tissues. Since epidermal cells do not contain chloroplasts, they are essentially colorless and facilitate the focusing of light to the active, photosynthetic mesophyll and palisade cells below. The stomatal pores present in the epidermis allow for gas exchange in photosynthesis and respiration, and are controlled by discrete signal transduction pathways that involve ABA, calcium, phosphatidic acid, and inositol-containing second messengers. These signal transduction components are thought to

eventually alter ion channel activities that allow the guard cells to increase turgor, thus opening the stomatal pore, or to decrease turgor, which results in stomatal closure. Thus, in addition to its role in seed dormancy, ABA is also considered the drought-sensing hormone as its signal transduction pathway can allow for stomatal closure, an important response to drought that conserves water lost through transpiration.

It is interesting to note that most leaves contain more stomatal pores on their abaxial surface than their adaxial surface. This location places them closer to the spongy mesophyll. Indeed, the mesophyll layer within the leaf is the major site of photosynthesis in the plant, and it contains two cell types in dicots: the spongy mesophyll and the palisade mesophyll cells. Both cell types are active in photosynthesis, yet they have different shapes. It is thought that the oblong shape of the palisade cells helps to further focus light on the spongy mesophyll cells. The gaps around spongy mesophyll are another adaptation that accommodates the oxygen generated from photosynthesis.

4.4.2. Leaf Development Patterns

Besides photosynthesis, there are several interesting developmental considerations for leaves. Leaf primordia first arise when a small group of cells on the outer edge of the SAM gain leaf identity. These leaf primordia mature into a leaf bud utilizing a marginal meristem to form the lamina or outer edge of the leaf, and a central meristem that gives rise to the vascular tissue. Leaf buds can remain dormant in plants such as trees. Cell division within the leaf bud occurs at the base of the primordia or leaf, which means that cells are pushed up toward the tip of the growing leaf. Along with cell division, cell expansion is a critical process that produces large increases in leaf size. In general, cell expansion starts after cell division has given rise to the main structure of the leaf. Thus, the younger the leaf, the more active it is in cell division. Almost all mutants defective in the production of leaves are also affected in the SAM, containing an under- or overcommitment to leaf primordia cells. Another interesting characteristic of leaves is their placement on the plant, which is called *phyllotaxy*. Leaves are initiated in a precise pattern as the shoot meristem grows, producing either alternate, opposite, tricussate (whorled), or spiral arrangements. In many species, the number and position of leaves, or modified leaves such as the spines of a pineapple fruit, follow the Fibonacci number series (1, 2, 3, 5, 8, 13,…). The venation pattern of leaves also varies with monocots containing parallel venation, while most dicot leaves have a reticulate pattern.

The shape of leaves is a very noticeable trait. Leaf shape is controlled by environmental and genetic programs as well as hormones. Some species such as tomato contain compound leaves, while others such as *Arabidopsis* contain simple, nonlobed leaves. Cell death within leaf primordia in plants such as philodendron produce "holes" in leaves. Some leaves such as pea also contain tendrils that function in "grasping" surrounding structures in the environment and facilitate directional growth. Corn leaves contain specialized domains called the *sheath, blade*, and *ligule*, which also facilitate growth by providing a way to change the position of the leaf surface, ensuring that photosynthetic tissues get maximal exposure to light.

Maize has been especially useful as a model plant to study leaf development. The *knotted* (*KN1*) gene, which is the related to the *shoot meristemless gene* (*stm*), mentioned in the discussion on SAM development, was first identified in corn mutants that contained knots of tissue on their leaves. These *KN1* mutants were defective in the normal regulation of the *KN1* gene, which would normally be confined to the apical meristem. Instead, *KN1* mutants contain *KN1* expression in the leaves, which results in an aberrant mass or knot of tissue. The corn *KN1* gene was ectopically expressed in transgenic tomato plants to investigate the role of this homeodomain transcription factor in dicot leaf development. The results were transgenic tomato plants containing an increase in leaf complexity. Recall that most tomato species contain compound leaves with several leaflets. Ectopic expression of the corn *KN1* gene caused a large increase in the number of leaflets per leaf, suggesting that in dicots, *KN1* can alter leaf complexity specification (Fleming 2006).

4.5. FLOWER DEVELOPMENT

4.5.1. Floral Evocation

Flowers are the flowering plant's most obvious and aesthetically pleasing organ. In general, all flowers are specified in a similar manner. For flower development to occur, vegetative meristems must first undergo a transition to produce the *inflorescence meristem*. These meristems are self-renewing and also give rise to the floral meristems that produce flowers. The term *floral evocation* refers to the process of inflorescence meristem commitment. This is controlled by many factors, including plant size, whether a cold season has passed (vernalization), environmental stress, and daylength. For example, short-day plants such as cocklebur and Christmas cactus require a minimum light period (<15 h) to flower. Only one inductive period of light is needed to block flowering in many short-day plants. In contrast, long-day plants, such as *Arabidopsis*, require a longer period of light (usually 12–16 h) to flower. *Arabidopsis* is also considered to be a long-day facultative plant, as it can flower in short-day conditions but will flower much faster if placed under long-day conditions. Daylength-neutral plants, such as tomatoes, are not as affected by the photoperiod.

After floral evocation has taken place, a plant can be moved to noninductive conditions and still flower. Many historical studies have suggested that a hormonal factor, termed *florigen*, is produced elsewhere in the plant, such as the leaves, and then stimulates floral evocation in the meristem. Trying to determine the identity of florigen has been a focus in plant biology for years because of its importance in agriculture. Flowers are the precursor of fruit; and if flowering can be controlled, plants can be manipulated to remain in a vegetative or flowering state. Accelerated flowering can lead to a much shorter growing season, which is advantageous for growers.

Not surprisingly, there are mutants defective in floral evocation, and their study helps us understand some of the molecular requirements for floral evocation. One such mutant, FLOWERING LOCUS T (FT), encodes a protein that is accepted to be a part of florigen. FT is a systemic inducer of flowering that is expressed in the companion cells of the phloem and is exported to the phloem sieve elements and then is transported to the shoot apex. FT-INTERACTING PROTEIN1 (FTIP1) is an essential ER-localized protein that interacts with FT and assists its transport from the phloem companion cells to the sieve elements. In the shoot, FT protein interacts with the transcription factors FD and LEAFY. FD and LEAFY are considered to be master switches that "turn on" expression of genes needed for flowering.

The *Constans* (*CO*) gene from *Arabidopsis* encodes a *zinc-finger transcription factor* whose mRNA levels rise and fall with a circadian rhythm. *CO* turns on a number of genes, including *FT*. Thus, CO protein accumulation, controlled by the circadian rhythm, triggers a cascade of events that results in flowering. More recent studies also indicate that increased CO protein expressed only in the leaves of transgenic plants can stimulate early flowering in *Arabidopsis*.

As mentioned earlier, LEAFY is a transcription factor involved in the switch from the inflorescence to floral meristem. *Leafy* mutants have a delay in floral meristem development and flowers are replaced by leaflike or flowerlike shoots, suggesting that the function of LEAFY is to promote floral meristem identity. Indeed, ectopic expression of LEAFY in transgenic aspen trees can speed up the flowering process in these trees, presumably by promoting floral meristem identity. Another important floral meristem mutant containing the opposite phenotype is the *terminal flower* (*tfl*) mutant. These *tfl* mutants flower early and have a determinate inflorescence which means that the inflorescence meristem is transformed into a terminal flower. Thus the function of the TFL protein is to promote inflorescence identity (Bernier and Perilleux 2005; Krizek and Fletcher 2005; Corbesier and Coupland 2006). Both FT and TFL encode proteins predicted to bind lipids. Indeed both FT and TFL1 have been recently shown to bind to the phospholipid phosphatidylcholine, which may allow them to associate with membranes in the plant cell (Nakamura et al. 2014, Hanzawa et al., unpublished observations).

4.5.2. Floral Organ Identity and the ABC Model

After floral evocation and development of a floral meristem committed to the process of flowering, the individual organs present in the flower develop. A flower consists of four concentric whorls containing flower organs that in most dicots like *Arabidopsis* are arranged this way: *sepals* (Se), *petals* (P), *stamens* (St), and *carpels* (C) (Fig. 4.8). Sometimes, one of the whorls is not well developed or is repeated (like the petals in a tea rose), or sometimes one whorl is dominant so that the rest of the organs are not noticeable. On closer inspection, however, one can usually distinguish the four types of organs.

The specification of floral organ identity begins during floral evocation, for example, when the LEAFY protein acts to turn on gene expression. We have learned the most about floral organ identity from *Arabidopsis* homeotic mutants. Floral homeotic mutants were isolated that contain a transformation of one organ into another. To understand these mutants and the resulting ABC model of floral organ identity genes, one must be familiar with the normal arrangement of organs in the *Arabidopsis* flower (Fig. 4.8). This flower contains an outer whorl of four green sepals: four white petals, four to five yellow stamens, and two fused carpels. *Agamous* (*ag*) mutants are homeotic mutants that are very striking and contain an outer whorl of sepals, followed by petals, and then sepals again. Comparison of *ag* flowers to wild-type flowers indicates that *ag* mutants have lost information required to make stamens and carpels in whorls 3 and 4, and have replaced this with petals and sepals, respectively. Mutants in *ag* also contain a reiteration of this pattern resulting in an indeterminate meristem and extra rows of petals and sepals. This finding indicates that AG function is required for whorls 3 and 4 (stamen and carpel) identity. In contrast to AG, the *apetela2* (*ap2*) mutants have sepals transformed into carpels in the first whorl, and petals transformed into stamens in the second whorl, followed by stamens and carpels in the next two whorls as usual. This indicates that AP2 is required for identity of whorls 1 and 2 (sepals and petals). Finally, two different mutants with the same phenotype, the *pistillata* (*pi)* or *apetela3* (*ap3*) mutants, contain a transformation of petals to sepals in the second whorl, and of stamens to carpels in the third whorl. This indicates that the PI and AP3 proteins function in identity of whorls 3 and 4.

Together, results from these homeotic mutants suggest that three separate types of genes (denoted A, B, and C) function in floral organ identity (Fig. 4.8). The A function is controlled by the AP2 gene product and must be required for both sepals and petals in whorls 1 and 2. AP3 and PI are gene

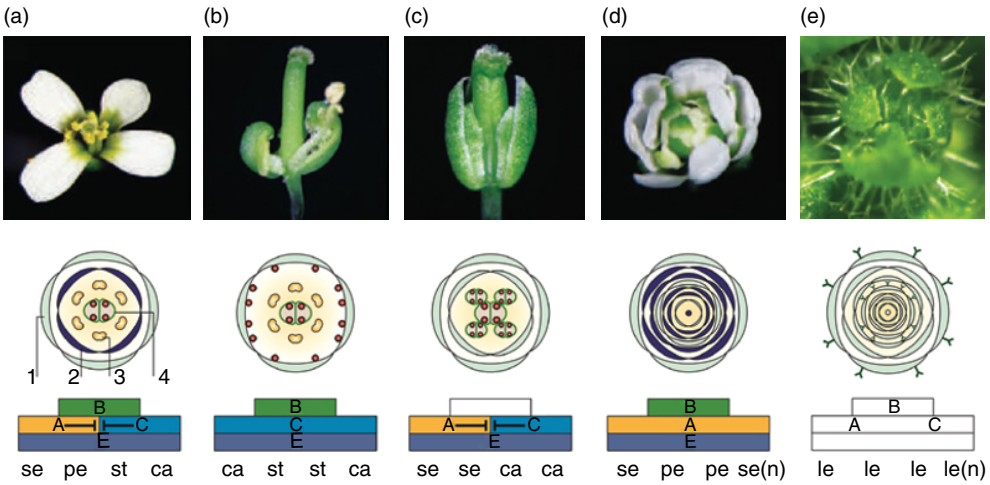


Figure 4.8. Flower development. *Arabidopsis* (a) wild-type, (b) *ap*2, (c) *pi*, (d) *ag*, and (e) *sep* flowers. Below each photo is a rendering of the ABC model as it functions in that flower. (From Krizek and Fletcher (2005). Reproduced with permission from Nature Publishing.) (*See insert for color representation of the figure.*)

products with a B function and are required in whorls two and three to help specify petals and stamens. Finally, the C function is controlled by AG, which helps specify whorl 3 (stamens) and whorl 4 (carpels). Important to this model is the antagonism of A/C function, such that if one is lost, the other expands its function into the two whorls where it would not normally function. Another caveat is that B function must necessarily be present in combination with either A or C to specify the petals and stamens. By drawing out each mutant's observed pattern, one can see that the mutant data "fit" exactly to this model.

This elegant model can also be used to predict the phenotypes of double and triple mutants, which, for the most part, verify the model. For example, if both A and B functions are lost, this model predicts that C function will expand to all four whorls, and that carpels should be present in each whorl. The resulting double mutant is found to contain a leaflike structure in whorl 1, carpelloid leaves in whorl 2, and carpels in whorls 3 and 4, a close approximation of what the model predicts. A triple mutant that has lost A, B, and C functions is predicted to contain no floral organ identity. The observed mutant is found to contain carpelloid leaves in each whorl, which suggests that the ground state of the flower is not totally vegetative (i.e., leaflike).

A new dimension to the ABC model has recently been discovered that involves a group of four genes, called *Sepellata* (*Sep*) genes, which are required to specify each whorl in addition to the ABC genes. Loss of this E function through a quadruple mutant lacking all four genes results in whorls of carpelloid leaves, similar to the mutant lacking ABC function.

Thus, our understanding of flower development starts with CO and LEAFY transcriptional function to begin the developmental program and results in the production of AP2, PI, AP3, AG, and SEP proteins. How do these proteins function to specify floral organs? AP3 and AG encode MADS box genes, a family of transcription factors expressed in yeast and plants that most likely function by turning on other, specific genes required to build a sepal, petal, stamen, or carpel. The ABC model predicts that expression of these genes should be confined to the specific whorls where they function. This prediction has been verified by observing the *in situ* mRNA expression patterns of the genes. For example, AP2 is expressed early in whorls 1 and 2.

It is important to note that several homeotic genes controlling floral development have been isolated from other plants, including *Antirrhinum* (snapdragon), supporting the importance of the ABC model in other species. For example, the *Antirrhinum deficiens* (*DefA*) gene probably functions similarly to AP3 from *Arabidopsis* (Krizek and Fletcher 2005).

4.6. HORMONE PHYSIOLOGY AND SIGNAL TRANSDUCTION

4.6.1. Seven Plant Hormones and Their Actions

Signal transduction is the cascade of events that allow a signal, usually from outside the cell, to be interpreted by the cell. Signal transduction cascades usually result in a final biological response, and often the response can be measured. Besides light and abiotic stress, the plant hormones are the major developmental and physiological signaling molecules in the plant. The seven major plant growth hormones are small molecules rather than proteins or peptides, and in some cases they are similar to certain animal cell hormones (Fig. 4.9). For example, brassinolide (Br) is a sterol, much like estrogen and testosterone, which function as sex hormones in animals. Br is critical for normal plant growth and development in plants, playing a role in stem elongation, leaf development, pollen tube growth, vascular differentiation, seed germination, photomorphogenesis, and stress responses.

Auxin, or indole 3-acetic acid, was the first plant hormone discovered and contains an indole ring much like the melatonin hormone of animals. Auxin is known to stimulate cell elongation and cell division, differentiation of vascular tissues, root initiation, and lateral root development. Auxin can also mediate the bending responses to light and gravity, and within the apical bud it suppresses the growth of lateral or axillary meristems. It can delay senescence, and interfere with leaf and fruit

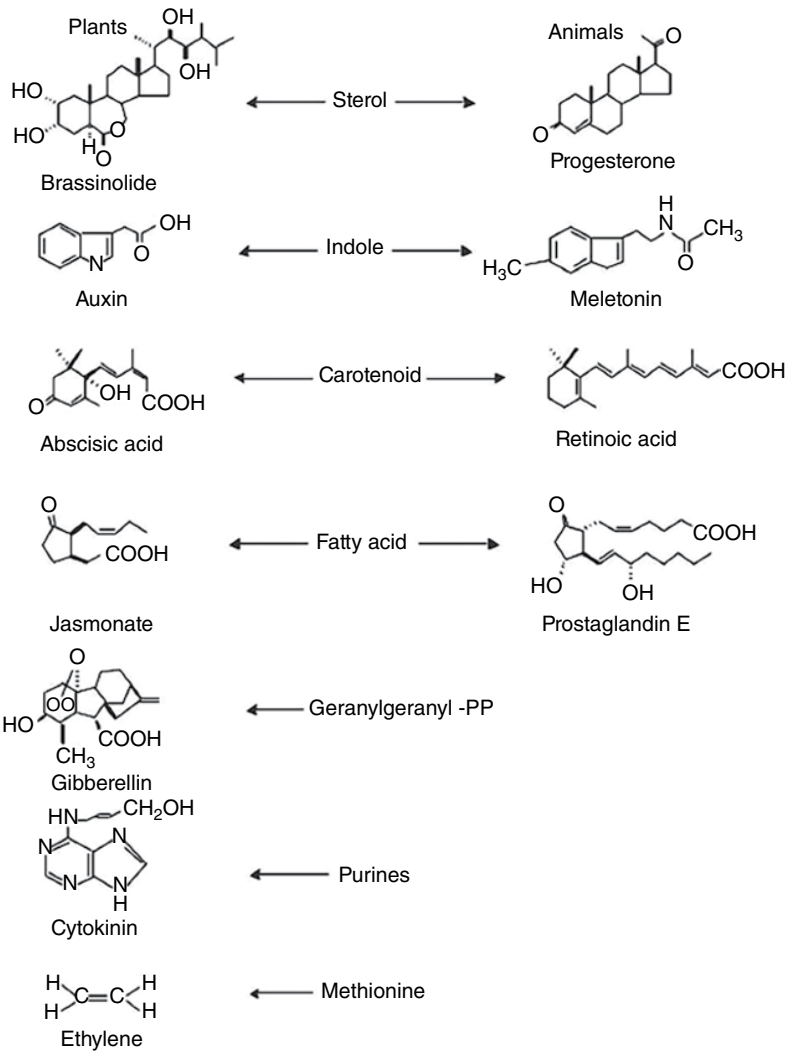


Figure 4.9. Plant hormones. Similarities between some plant and animal hormones. (From Chow and McCourt (2006). Reproduced with permission from Cold Spring Harbor Laboratory Press.)

abscission. It can induce fruit setting and delay ripening in some fruits. It can also stimulate the production of another plant hormone, *ethylene*.

Cytokinin is generally considered the second most important plant growth-regulating hormone, following auxin. Cytokinin is similar to adenine and was first discovered in 1941 as the active component in coconut milk that promoted growth of plant cells in tissue culture. Cytokinin can promote cell division and shoot growth and can delay senescence.

ABA was first identified in a search for an abscission-promoting hormone. This is not the function of ABA; and as noted earlier, it functions in promoting dormancy and in sensing drought and other stresses. ABA is derived from mevalonic acid and carotenoids and is thus similar in structure to the developmental factor from animals called *retinoic acid*. Transport of ABA can occur in the vascular tissues. ABA stimulates closure of the stomatal pore and can inhibit shoot growth. In seeds, it promotes dormancy and stimulates the production of seed storage proteins. It is mostly antagonistic to GA and can inhibit the response of grains to GA. ABA is also involved in inducing gene transcription in response to wounding, which may explain why it has a role in the pathogen defense response.

Jasmonic acid (JA) is a fatty-acid-derived plant hormone that is similar in overall structure to physiologically active small molecules from animals called *prostaglandins*. In plants, JA is firmly associated with pathogen defense pathways. For example, it has been documented that the physical stimuli of certain insects can trigger the synthesis of JA, which then functions to increase expression of genes involved in defending the plant, such as the *pathogenesis-related 1* (*Pr1*) gene. Microbial and viral pathogens can also trigger JA synthesis, thus the study of JA-mediated events in the plant cell are of interest to plant pathologists who wish to engineer transgenic plants that are disease-resistant.

GA and ethylene are two plant hormones with no similar molecular counterparts in other eukaryotic organisms. GA was first discovered from fungi that can stimulate plant cell elongation and cause significant and "leggy" growth of rice plants. GA is a series of 136 diterpene compounds that contain 19 or 20 carbons in four- or five-ring systems. These are named for the order in which they were discovered (GA1, GA2, etc.). The other functions of GA, as mentioned previously, are in general antagonistic to the actions of ABA. For example, ABA promotes seed dormancy, while GA is required in most cases to break seed dormancy. The actions of GA on barley germination have been well studied where it has been shown that GA promotes expression of the *α-amylase* genes required to break down starch in barley aleurone, an important process in the grain-malting business. GA also plays a prominent role in stimulating flower development under long days.

Ethylene, a hydrocarbon gas, is a very simple molecule that is best known for its stimulation of fruit ripening and promotion of the seedling triple response. Indeed, people of ancient cultures understood the actions of ethylene and could burn incense in a closed room to stimulate fruit ripening. The triple response of seedlings is a specific developmental program wherein an apical hook forms in the shoot, and the root becomes thicker. These adaptations may increase survival under certain conditions. In addition, ethylene can stimulate the release of dormancy, adventitious root formation, flower opening, and flower and leaf senescence.

4.6.2. Plant Hormone Signal Transduction

The first eukaryotic *signal transduction* pathways to be characterized were the peptide growth hormone pathways of animal cells. This most likely resulted from the important discovery that animal oncogenes sometimes encoded altered growth factors, growth factor receptors, or other signal transduction components that regulate cell growth. Often, these components function in a signaling cascade in which sequential protein phosphorylation plays an important role in activating proteins. A paradigm signal transduction pathway is shown in Figure 4.10 to facilitate understanding of how signal transduction works. Hormones on the outside of a cell can be perceived, either by receptors that span the plasma membrane, or by proteins on the inside of the cell. After initial stimulation by the hormone, information can be relayed by a signal cascade of small molecules and/or proteins. Often, this signal cascade impacts a transcription factor in the nucleus, where activation can stimulate new gene expression programs. The resulting gene expression results in the production of new proteins that can function in the final biological responses to the hormone.

Because plant hormones are small molecules rather than proteins, and because the plant cell wall encloses the plasma membrane, plant hormone signal transduction pathways are sometimes significantly different from those of animals. However, it is important to keep in mind that most of the individual components of plant signal transduction pathways have similar counterparts in other eukaryotes. Plant receptors linked to plant hormone action were not discovered until the 1990s. The accelerated pace of experimentation that followed resulted in major paradigms of plant hormone signal transduction (Gibson 2004; Chow and McCourt 2006).

4.6.2.1. Auxin and GA Signaling. Plants use auxin to regulate important aspects of growth and development. When auxin acts to promote cell division and growth, it does so mainly by increasing the expression of genes that encode required proteins for these processes (as shown in Fig. 4.10).

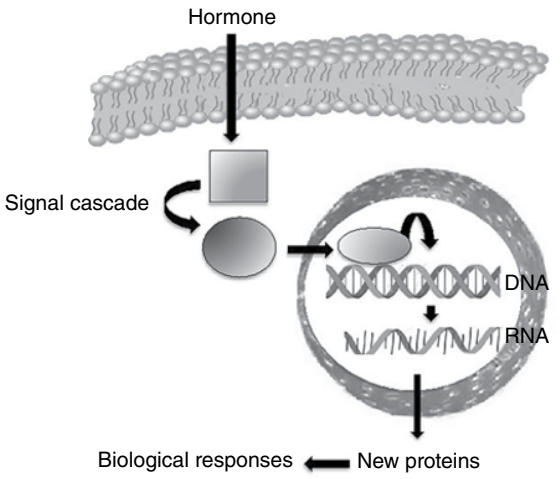


Figure 4.10. A paradigm plant hormone signal transduction pathway. Hormone on the outside of a plant cell may be perceived by proteins present at or near the plasma membrane. Alternatively, the hormone may be transported across the plasma membrane. Signal cascade proteins are then activated. Once activated, these proteins can transmit signaling information (arrows) to the interior of the cell. Many signal transduction pathways converge on the stimulation of gene expression within the nucleus that results in the production of new proteins in the cytoplasm that can affect specific biological responses.

Thus, researchers have sought to understand the steps between auxin perception and the final gene expression regulation. We now know that auxin signaling involves ubiquitin-mediated protein turnover as way to control transcription of genes that allow the plant to effect a response to auxin. Molecular studies revealed the first players in auxin signaling as a group of genes encoding the IAA/AUX proteins whose expression is rapidly upregulated in response to auxin within minutes. Most of the IAA/AUX proteins are nuclear-localized and have a very short half-life. They can form heterodimers with the *auxin-response factor* transcriptional regulators (ARFs), and then bind to a 6 bp (6-base-pair) auxin-responsive element (ARE) present in the promoters of auxin regulated genes. Further studies revealed that ARF:ARF homodimers were responsible for activation of gene expression in response to auxin, while ARF:AUX/IAA heterodimers blocked transcriptional activation (Quint and Gray, 2006).

Genetic mutants that failed to respond to auxin in seedling growth assays identified genes that are required for some of the auxin signal cascade. These genes, which include the *axr1* and *tir1* genes, encode proteins that function in the ubiquitin-mediated protein turnover pathway in eukaryotes. The proteasome is a large, macromolecular structure that functions to degrade proteins within the cell. Proteins destined for the 26S proteasome are modified by the addition of ubiquitin, itself a small protein (76 amino acids). Thus, auxin signaling requires a functioning 26S proteasome and enzymes necessary to add ubiquitin to target proteins.

It has been shown that the F-box protein encoded by TIR1 becomes physically associated with auxin, plus a small inositol hexakisphosphate ($InsP_6$) signaling molecule, and the IAA/AUX proteins. Thus TIR1 is considered to function as an auxin receptor. After binding to auxin, TIR1 stimulates the proteasome to specifically degrade the bound IAA/AUX proteins. Once the IAA/AUX proteins are degraded, ARF:ARF homodimers form, bind the AREs in promoters of auxin-responsive cells, and stimulate the transcription of these genes. In this way, auxin can stimulate expression of genes required to carry out its physiological effects.

Auxin perception and transcriptional activation are mediated by the degradation of Aux/IAA repressor proteins. Degradation of Aux/IAAs relieves repression on ARFs, which bind DNA sequences called auxin-response elements (AuxREs). In most higher plant genomes, multiple

paralogs exist for each part of the auxin nuclear signaling pathway, including the TIR1 receptors. This potential combinatorial diversity in signaling pathways likely contributes to the myriad of context-specific responses to auxin (Schaller et al. 2015).

Interestingly, proteasomal degradation of a transcriptional repressor may be a common theme in plant hormone signal transduction pathways. The DELLA protein represses GA-regulated genes and is degraded by the proteasome after GA addition to plant cells. Thus, GA signaling may share the same general regulation in stimulating expression of genes required for the physiological responses to GA. The GA signal transduction pathway also has an identified receptor, GID1, which functions to bind GA and act in the initial step in GA perception. GID1 is a nuclear soluble protein that is homologous to the animal hormone-sensitive lipases and GID1 binds to different GAs with saturable kinetics. This last fact is an important test that helps support the idea that GID1 directly and specifically interacts with GA (Pimenta-Lange and Lange 2006). The other required component for DELLA degradation by the proteasome is the F-box proteins SLEEPY1 (SLY1) and SNEEZY (SNZ) in *Arabidopsis* and GIBBERELLIN-INSENSITIVE DWARF2 (GID2) in rice. The formation of the GA–GID1–DELLA complex enhances recognition between SLY1/GID2 leading to promotion of the ubiquitylation and subsequent destruction of DELLAs by the proteasome, thereby relieving the growth-restraining effects of DELLA (Daviere and Achard, 2013).

4.6.2.2. Cytokinin and Ethylene Signaling. Plant cells utilize elements of the *two-component signaling pathways* in their responses to cytokinin and ethylene. The two-component systems function in microbes, yeast, and plants to convey signals between a histidine kinase receiver and a phosphorylated response regulator (RR). These two components are joined by an intermediate in plant cells termed the *phosphorelay intermediate*. Both cytokinin and ethylene have been shown to bind to specific histidine kinases that function as receptors, contained in the plasma membrane. This binding is thought to stimulate a phosphorylation cascade wherein the activated histidine kinase phosphorylates an intermediate protein, which then phosphorylates a specific aspartate residue on the RR. The RR then acts to stimulate downstream functions that in plants involves transcriptional controls.

Arabidopsis has three cytokinin receptors (AHK2, AHK3, and CRE1/WOL/AHK4) that contain a conserved cytokinin binding domain, a histidine kinase domain, and a receiver domain. Thus the AHK receptors play partially redundant roles in cytokinin signaling. The AHKs are primarily located in the ER membrane in plants, suggesting that the site of cytokinin binding is in the lumen of the ER. After binding cytokinin, the *phosphorelay intermediates*, called AHPs are phosphorylated. AHPs act as intermediates in the transfer of the phosphate to the downstream RRs. There are five functional AHPs in *Arabidopsis* and these can be actively transported in and out of the nucleus, independent of their phosphorylation status. The downstream targets of the AHPs, the RRs, fall into two main classes called type-A and type-B RRs. The type-B RRs are transcription factors that contain a specific DNA-binding domain called "Myb." There are multiple type-B RRs in *Arabidopsis* that act as partially redundant, positive elements that control the transcriptional response to cytokinin. Interestingly, the protein levels of some type-B RRs are regulated by F-box proteins and the proteasome. The type-A RRs lack a DNA-binding domain, negatively regulate cytokinin signaling, and are rapidly induced by cytokinin. Thus, the type-A RRs function as negative feedback regulators in cytokinin signaling (Schaller et al. 2015).

In ethylene signaling, five ER-localized proteins (ETR1, ETR2, EIN4, ERS1, and ERS2) bind to ethylene and are modified two component system histidine kinase receptors. In a pathway that seems reminiscent of auxin and GA signaling, these receptors signal via signal cascade proteins triggering proteasomal degradation of the EIN3 transcriptional activator. One key difference is that these ethylene receptors are active in signaling in the absence of ethylene. The intermediary proteins activated in the absence of ethylene are a Raf-like Ser/Thr protein kinase CTR1 (CONSTITUTIVE TRIPLE RESPONSE 1). Active CTR1 directly phosphorylates an ER protein named EIN2, which becomes degraded by the proteasome in its phosphorylated state. The degradation of EIN2 leads to

EIN3 degradation in the nucleus. Since EIN3 is a transcriptional activator of ethylene-responsive genes, the absence of ethylene leads to suppression of expression of ethylene-responsive genes.

Interestingly, in the presence of ethylene, the receptors are inactivated. CTR1 does not phosphorylate EIN2, and EIN2 becomes cleaved into two protein fragments. The C-terminal fragment (CEND) then exits the ER, and is imported into the nucleus where it functions to help stabilize EIN3. Stabilized EIN3 activates expression of the ERF transcription factors and other ethylene responsive genes (Yang et al. 2015).

4.6.2.3. Brassinosteroid Signal Transduction. The brassinosteriod, brassinolide (Br), is a more recently discovered plant hormone. As was carried out for the other plant hormones, genetic mutant screens were performed to find Br-insensitive mutants. The *BRI* gene was identified and shown to be required for seedling responses to exogenously added Br. The BRI protein encodes a leucine-rich repeat (LRR)-containing serine/threonine protein kinase, and has two homologues, BRL1 and BRL3, that also can function as Br receptors. The presence of an LRR domain is important since LRR signaling kinases are abundant in animal cells and often serve as receptors for animal peptide hormones such as insulin. The BRI protein is predicted to span the plant cell plasma membrane, making the LRR domain accessible to the outside of the plant cell, with the kinase domain contained on the interior of the cell. This arrangement led to an integral domain-swapping experiment between the BRI protein and the XA1 protein that confers resistance to rice blast fungus. Researchers produced transgenic plants containing the outside LRR domain from BR1 and the interior kinase portion of XA1. The resulting plants could be stimulated with Br to turn on disease resistance pathways, cleverly showing that each part of these receptors is specific and can function when swapped.

When Br ligand binds to the BRI receptor, Br is proposed to serve as a "glue" that promotes phosphorylation of the negative regulator BRASSINSOSTEROID KINASE INHIBITOR1 (BKI1), and subsequent association with the BRI1-ASSOCIATED KINASE 1 (BAK1). As a result of BRI binding to BAK1, both undergo reciprocal transphosphorylation, resulting in an enhanced signaling output via phosphorylation of other protein kinases and phosphatases. This enables the subsequent activation of BRASSINAZOLE RESISTANT1 (BZR1) and a homologous transcription factor BRI1-EMS-SUPPRESSOR1 (BES1). Together, these transcription factors regulate the expression of hundreds of Br-regulated genes that impact plant growth and development (Singh and Savaldi-Goldstein, 2015). In this way, the Br signal transduction pathway is similar to other signaling pathways we have examined, with a final nuclear transcriptional output required for the final biological response to Br. It is interesting to note that there are over 170 genes in *Arabidopsis* predicted to encode LRR kinases that may function in plant signal transduction. In addition, there are several other putative receptor-like kinases that do not contain LRR domains, but also may function in signaling.

4.6.2.4. JA Signal Transduction. JA is a defense hormone whose isoleucine (Ile)-conjugated form binds to a protein called CORONATINE-INSENSITIVE 1 (COI1). Structurally, COI1 encodes an F-box that is closely related to the TIR1 auxin receptor. Like TIR1, COI1 binds to so-called repressor proteins named JASMONATE-ZIM DOMAIN (JAZ) repressor proteins, which normally act to repress JA-regulated genes. And not surprisingly, binding of JA-Ile to COI and JAZ proteins requires a small molecule, inositol pentakisphosphate ($InsP_5$), leading to proteasomal degradation of the JAZ repressors. When the JAZ repressors are degraded, the MYC2 transcription factor can bind to promoters of JA-regulated genes and promote expression. Thus the JA signaling pathway shares many characteristics with auxin signaling (Pérez and Goossens 2013). This brings up a critical point: plant hormone signaling pathways undergo "cross-talk", which is a type of connections resulting in co- or cross-regulation of two or more pathways. While outside the scope of this chapter, good examples of cross-talk are the ARF auxin transcription factors (ARF6 and ARF8) regulating JA biosynthesis in flowers, auxin inducing JAZ1 expression, and MYC2 mediating the repression of genes involved in auxin-induced regulation in the root meristem. In fact, understanding the complex interplay of hormone cross-talk in plants is now an exciting area of research.

4.7. CONCLUSIONS

Even though plants do not have elaborate body plans and the number of specialized cells as organs compared with animals, the developmental programs of plants are no less elaborate. A number of crucial plant growth regulators or hormones are required for proper plant growth. We will see in Chapter 5 how plant biotechnologists can alter hormone type and concentration to manipulate cells in Petri dishes, a general requirement for plant genetic transformation.

LIFE BOX 4.1. DEBORAH DELMER

Deborah Delmer, Professor Emeritus UC Davis; Rockefeller Foundation (retired); Winner of the ACS Anselme Payen Award; Member of the National Academy of Sciences.



Deborah Delmer. Courtesy of Deborah Delmer.

I must confess that there was something rather haphazard in the path I took to become a serious scientist. A major positive influence was my father—a small-town country doctor in Indiana, who had a passion for his work that certainly impressed me. Ours was a family in which Mom and my brother had a very close alliance, while the same was true for me and my father. And so, I suppose it was natural that he hoped very much that I would follow in his path—which meant enrolling at Indiana University and pursuing a career in medicine. I personally also found this attractive, but my boyfriend had other ideas—that I should study rather to be a nurse—a career that should be more suitable for what he hoped would be my main calling in life—his wife and mother to his children. As it turned out, I pleased neither of them. During my first week at Indiana University, I hoped to still my own confusion by talks with faculty at orientation day. I started in alphabetical order and, within half an hour, had signed up for anthropology as a major. But then I wandered on to "B," and there was this handsome young professor who had a crowd of students enthralled by his passionate advocacy of the field of bacteriology. I joined the crowd and asked him, "Could I major in this as a Pre-Med?" "Yes" "Isn't it dangerous to work with bacteria?" "No, it's FUN!!!!." And I was hooked. Major A changed to Major B and I never looked back. For some freaky reason relating to the fact that I was an honors major, I had the Chairman of the English Department as my advisor freshman year, and he urged me to go for a BA degree instead of a BSc—and this turned out also to be quite lucky—in addition to science, I took honors courses in creative writing, advanced English literature, and also many extra courses in Russian and spent a summer in Russia back when the cold war was really cold. It's true that I lived my life with a secret fear all my life that I never was a strong in math and chemistry as my other colleagues. Yet, I really hate the specialization we impose on our science majors now—and have no regrets at having such an enriching undergraduate experience.

I loved bacteriology—I think as much for the terrific faculty as for the discipline—and, to my father's disappointment, I decided that graduate school was a more appealing choice than medical school. Escaping the

boyfriend meant going away as far as possible from Indiana for graduate school—and so I chose the field of marine microbiology at the Scripps Institution of Oceanography in California—but soon found that I got seasick easily. Again, on a random choice, I moved sideways to the new biology department on the new campus of UC San Diego. Again, fate played a role, and I was given a rotation with Carlos Miller who was on sabbatical at UCSD. Carlos had a key role in the discovery of the plant hormone cytokinin, and was a lovely gentle fellow who had much patience with this student who had never studied botany because of all those English and Russian courses. But he convinced me to stay with plants, and I ended up doing my thesis work characterizing the pathway for tryptophan biosynthesis in plants. By then, I was married to a graduate student in astrophysics who was offered a great postdoctoral opportunity at the University of Colorado. Wanting to stay with plants, I arranged my own postdoc at Colorado with an up-and-coming young fellow named Peter Albersheim who was just beginning his groundbreaking work on the structure of the plant cell wall. It was this focus that was to set me on my own career path that focused for the rest of my academic career on the study of plant cell walls.

Although Peter had concentrated on cell wall structure, I felt more inclined to enzymology and decided to tackle a major unanswered question that occupied me for the rest of my career—the mechanism of biosynthesis of the world's most abundant organic compound—cellulose. By now, I was a young faculty member at Michigan State University, and I chose the cotton fiber as a model system because it was a veritable factory for cellulose. We struggled without success trying to identify an enzyme system that could make cellulose—but here is a lesson for the young. While still maintaining our focus on the key issues, I also knew that one has to show productivity—and so I initiated some other projects that were "easier" to succeed with—the first demonstration of the role of lipid intermediates in plant glycoprotein synthesis, the pore size of plant cell walls, insights into the biosynthesis of callose, and a rather comprehensive characterization of cotton fiber development.

Again, fate intervened when, for personal reasons, I relocated to The Hebrew University in Jerusalem. There we continued to focus on cellulose biosynthesis with two of my "favorite" projects—the finding that sucrose synthase—a key enzyme in synthesis of the precursor to cellulose, UDP-glc, had a plasma membrane-associated as well as the well-known soluble form, and the discovery that cells adapted to growth on an inhibitor of cellulose synthesis could survive with almost no cellulose in their walls—the latter showing just how adaptable plants can be when challenged. But the enzyme cellulose synthase still remained elusive. And here we can learn another lesson—don't be afraid to collaborate and delve into new areas of science. In order to get more comfortable working in molecular biology, I arranged a sabbatical with Dave Stalker at Calgene, Inc., in Davis, California. Dave's group was interested in getting more good cotton fiber-specific promoters, and we were interested to try to identify the gene for cellulose synthase—so we combined forces using our own cotton fiber cDNA library—Dave got his promoters and together we identified for the first time two sequences that encoded proteins that had all the domains expected for a cellulose synthase (plus a few more interesting domains!) and was highly expressed just at the time fibers underwent a 100-fold increase in cellulose synthesis as they initiated secondary wall synthesis. Discovery of these genes allowed the *Arabidopsis* gurus to find similar genes and show that when disrupted they did indeed lead to loss of ability to synthesize cellulose synthesis. From there, the field now has been joined by a healthy number of young new faces, and new discoveries about the process seem to emerge every month. We too found that *Arabidopsis* had many advantages and used it to advantage once I relocated my lab to UC Davis. Yet the power of being able to combine my old skills in biochemistry with the new skills in molecular biology I think has proved to be a very important aspect of my work.

Finally, my dad—who loved medicine because it combined good science with helping people—would be proud of me at last. At age 60, I made the unusual choice to retire from academia and work for the Rockefeller Foundation where I spent 5 years developing a portfolio of grants that built capacity in biotechnology in the developing world—especially sub-Saharan Africa—and supported projects that aimed to demonstrate that

biotechnology can offer solutions to at least some problems that breeders find intractable. I've enjoyed this new challenge immensely—and now have another new one—retirement! But I continue to consult on issues of international agriculture and, in a twist, have also found my knowledge of cellulose synthesis again valuable to those working on biofuels. So it's been an interesting life—and it's not over yet!

LIFE BOX 4.2. NATASHA RAIKHEL

Natasha Raikhel, University Distinguished Professor, Ernst and Helen Leibacher Endowed Chair, University of California Riverside; Member of the National Academy of Sciences.



Natasha Raikhel. Courtesy of Natasha Raikhel.

I originated from and grew up in the Soviet Union. I immigrated with my husband and first-born son to Athens, Georgia, in 1978 (my second son was born in Athens, Georgia) with a personal fortune of only $25. I remember feeling somewhat lost and wondered how I could and would ever make the language, scientific and social transitions required of me. I did not realize at the time that I was lucky in many ways and that fortune had favored me.

I knew only one American scientist when I first arrived, but I encountered many helpful people that were critical to my survival. I also entered a social context within academia that differed in several important ways from the system I left behind. The American academic system is characterized by greater diversity and openness of thought and a healthy atmosphere of competition that drives one to take intellectual risks and achieve more. At its best, this environment also leads to a constant renewal of possibility, a wealth of new ideas and a rich milieu of thoughtful exchange that fosters both collective and individual progress. In America, I found a place where prestige and intellectual and economic rewards were all reasonable potential goals. Although I did not find the streets paved with gold, I actually found the far greater treasure of opportunity.

What I achieved was also due to timing. I am a product of this age of molecular biology, and its corollary age of rapidly expanding knowledge bases and burgeoning information systems that our technological growth has made possible. This lucky moment in history has allowed all of us the privilege to be pioneers of new and fascinating frontiers. When I came to this country, molecular approaches in plant biology were just beginning. I did not have to catch up, because I learned along with many people who were also just beginners in molecular biology. So, once again, I was lucky with good timing.

I am a cell biologist working with plants. I am fascinated by plants: we live on this planet because of plants, and I want to unlock secrets of plant cell biology. In my laboratory, we are using a model plant: Arabidopsis. This plant has a small genome and has been sequenced with many of its proteins identified. It is, therefore, a very convenient model organism for studying processes that are important in all plants including crop plants. My group has worked on the trafficking of molecules through the cell's vesicles and vacuoles, and we are interested in the synthesis of the cell wall in plants. A cell contains compartments called "organelles."

Compartments in cells are necessary to isolate and secure a large number of molecules that play an individual role(s) in various functions of the cell. In order for cells to function properly, molecules have to be produced and delivered to their proper destinations within the cell. Because plants are immobile and cannot run, they have to be

very versatile in their ability to respond to environmental stresses and survive. Therefore, plant cells have evolved a highly complex organization of functions to sustain life. The failure of any of these functions could poison other dynamic processes occurring within the intracellular environment and actually cause the destruction of the entire cell. Alternatively, improving the success rate of sending novel proteins and carbohydrates to desired parts of the cell can result in the improved nutritional value of crops and increased biomass production.

We live in an era of unprecedented biological discovery. Technologies to sequence entire genetic codes have yielded a wealth of data that require a focused interdisciplinary approach to assimilate and exploit this information. Once we understand the functions of all gene products (proteins), how they interact, and how pathways in the cell interact, we can really start to answer questions about how cells function and how the whole plant works. We call this new science "systems biology." The essence of systems biology is to model organisms and predict how various pathways in the organism interact when one pathway is affected. This requires the infusion of plant biology with disciplines such as mathematics, statistics, informatics, chemistry, and engineering.

It is very important that the new generation of plant biologists have multidisciplinary experience and training. I think that the community of *Arabidopsis* researchers will make the systems approach work because they are exemplary forward thinkers, effective trainers, and extremely open in sharing knowledge and tools. I hope that many talented young students are drawn to plant biology. Our field allows young people to reach for the stars and grow to the best of their potential.

I have tried, as I built a career as an American scientist, to foster and mentor those who will carry our field on into the future, to be persistent in the pursuit of worthy goals, and to change and learn new things when this is necessary. Although the research in my own group is extremely important to me, I have realized that I have experience that enables me to do more for the scientific community. Lately, my career has shifted somewhat from building a personal reputation towards accepting the responsibility of leadership within our field. But leadership does not occur in isolation. We all lead and follow within a group, hopefully as a team. In his essay, "Tradition and Individual Talent," the poet T. S. Eliot says that no artist has his complete meaning alone. I would expand that thought to include today's scientist, who also cannot have his or her complete meaning alone. It is the American context, which, at its best, celebrates diversity, the acceptance of new ideas and the ever-present possibility to start again and create a wonderful life.

LIFE BOX 4.3. BRENDA S.J. WINKEL

Brenda S.J. Winkel, Professor and Head, Department of Biological Sciences, Virginia Tech.



Brenda Winkel on research leave with the family in the Netherlands in 2009. Courtesy of Brenda Winkel.

It is hard to imagine that I could have planned my career to unfold as it has. The eldest child of Dutch immigrants, my mother's ambitions for her children were high, and at a young age there emerged a plan that I would someday become a physician. That was not to be, however, and looking back, it seems my life as a scientist has been driven, in large measure, by a regular series of serendipitous events, encounters, and discoveries. Growing up in a family with a love of plants, gardening, and the outdoors was also a great setup for latching onto the fast-emerging field of plant molecular biology in the mid-1980s.

One of the most memorable early "jolts" to my career plans was my first exposure to

a research seminar, as a senior chemistry undergraduate. This experience, together with what turned out to be fortuitous confusion surrounding my state residency for medical school, led instead to pursuit of a master's degree in biochemistry. A subsequent segue into industry, working at a new biotechnology division at Pall Corporation on Long Island, focused my sights on graduate school again, as I discovered that the PhDs were the ones with the really fun jobs! This realization, together with reports of the first transgenic plants in 1983, set me squarely on the path to a Ph.D. in plant molecular genetics.

By good fortune, I landed in Rich Meagher's lab at the University of Georgia in Athens in 1986. Rich was a molecular biology pioneer who had cloned several of the first plant genes and kept bringing new technologies and ideas to his lab, including plant transformation and PCR while I was there. It was also through Rich that I first met Tom Gerats, visiting from the Free University in Amsterdam, who replicated his remarkable collection of petunia anthocyanin mutants in the Athens greenhouses. Just as I was thinking about postdocs, Rich introduced us to *Arabidopsis*, which we grew in trays in a windowsill. That was the new frontier in plant biology! More good fortune, and Rich's connections from his own postdoc at UC-San Francisco, landed me a position with Howard Goodman at Massachusetts General Hospital/Harvard Medical School. There Howard was helping put *Arabidopsis*, quite literally, on the map, and growing plants in greenhouses on the roof of a new research building, 12 stories above the hospital grounds and overlooking the historic MGH "ether dome."

Howard offered the very enviable, if somewhat daunting, opportunity to develop my own research project, as long as it involved *Arabidopsis*. Serendipity was to enter the picture again. It turned out that a fellow postdoc was using one of Maarten Koorneef's transparent testa (tt) mutants as a marker for positional cloning. At the time, only one other lab was working on *Arabidopsis* flavonoid genes; one floor below, Rhonda Feinbaum in the Ausubel lab had isolated chalcone synthase for their work on host–pathogen interactions. So I set out to clone more flavonoid genes from *Arabidopsis*. After 9 painful months, there had been little discernable progress. Then one evening at a Boston *Arabidopsis* Network meeting, Brian Keith from the Fink lab at MIT explained how he'd used degenerate oligonucleotides to amplify target genes from genomic DNA. Within a couple of weeks I had pulled out first the chalcone isomerase (CHI)– and then the dihydroflavonol reductase (DFR) genes—and my postdoc (and career in flavonoids) was back on track! Not that everything was perfectly smooth from then on. As an example, there was the memorable occasion when I dropped a genomic DNA gel on the floor, shattering it into pieces and delaying, by several days, the moment I stood watching the x-ray developer spit out the image of a Southern blot that clinched the identification of tt3 as the DFR locus.

It was while interviewing for faculty jobs in my third year in Howard's lab that perhaps the most important "jolt" occurred, one that finally crystallized all the chance happenings into a clear and exciting new way forward. Three floors below us, Roger Brent's group had just developed one of the first yeast-two hybrid systems, which a fellow postdoc suggested I try on a proteasome subunit gene I'd stumbled on. Instead, inspired by the work of Geza Hrazdina at Cornell on flavonoid enzyme assemblies, I used it to generate the first evidence of direct interactions among CHS, CHI, and DFR. I recall telling my department-head-to-be on a return visit to Virginia Tech in Blacksburg that I had some very exciting new results. These became the basis for my first successful grant proposal, funded within my first year on the faculty (those were the days!), and research that continues in my lab to this day.

Over the past 22 years in Blacksburg, I have had the great good fortune to work with many fabulous students and postdocs, every one of whom brought their own ideas, biases, and passions to our research effort. There have been many more defining moments and "jolts" that included the surprising discovery of flavonoid enzymes in the nucleus and, more recently, evidence that links this puzzling finding to an apparent epigenetic role for chalcone synthase in plant defense. I have also had the privilege of working with chemistry colleague, Karen Brewer, and her group, from the time we met fortuitously at new faculty orientation until her untimely death last fall. This work, on the development of multimetallic

anticancer agents, has brought another completely unexpected and very rewarding dimension to my scientific life.

A "family" sabbatical (my husband is also on the faculty at Virginia Tech) in 2009, to Wageningen University in the Netherlands, was a long-overdue opportunity to reflect and recharge. The trip took me back to my ancestral and scientific roots, meeting up again with Tom Gerats and Maarten Koornneef as well as numerous aunts and cousins. It was also the prelude to another dramatic shift in my career that took place not long after our return to Blacksburg. Although I'd always been certain that administration was not for me, it suddenly seemed to be the right time—my department needed a new head and I was ready for a new challenge. Although my research has certainly been impacted by this latest jolt, it continues to be shored up by indulgent colleagues at Virginia Tech and at Wake Forest University, while being department head has pushed me to grow in other ways. Of course, it's still not clear what lies around the next turn, in research or my career, but that's really the fun of it all.

REFERENCES

Berger F, Grini PE, Schnittger A (2006): Endosperm: an integrator of seed growth and development. *Curr Opin Plant Biol* **9**:664–670.

Bernier G, Perilleux C (2005): A physiological overview of the genetics of flowering time control. *Plant Biotechnol J* **3**:3–16.

Boavida LC, Becker JD, Feijo JA (2005): The making of gametes in higher plants. *Int J Dev Biol* **49**:595–614.

Byrne ME (2006): Shoot meristem function and leaf polarity: the role of class III HD-ZIP genes. *PLoS Genet* **2**:e89.

Campilho A, Garcia B, Toorn HV, Wijk HV, Scheres B (2006): Time-lapse analysis of stem-cell divisions in the *Arabidopsis thaliana* root meristem. *Plant J* **48**:619–627.

Chow B, McCourt P (2006): Plant hormone receptors: perception is everything. *Genes Dev* **20**:1998–2008.

Corbesier L, Coupland G (2006): The quest for florigen: a review of recent progress. *J Exp Bot* **57**:3395–3403.

Costa S, Dolan L (2000): Development of the root pole and cell patterning in *Arabidopsis* roots. *Curr Opin Genet Dev* **10**:405–409.

Daviere JM, Achard P (2013): Gibberellin signaling in plants. *Development* **140**: 1147–1151.

Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, Hahn MG, Feldmann KA, Benfey PN (1996): The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell* **86**:423–433.

Fleming AJ (2006): Leaf initiation: The integration of growth and cell division. *Plant Mol Biol* **60**:905–914.

Gibson SI (2004): Sugar and phytohormone response pathways: navigating a signalling network. *J Exp Bot* **55**:253–264.

Koornneef M, Bentsink L, Hilhorst H (2002): Seed dormancy and germination. *Curr Opin Plant Biol* **5**:33–36.

Krizek BA, Fletcher JC (2005): Molecular mechanisms of flower development: an armchair guide. *Nat Rev Genet* **6**:688–698.

Lenhard M, Laux T (1999): Shoot meristem formation and maintenance. *Curr Opin Plant Biol* **2**:44–50.

McCormick S (2004): Control of male gametophyte development. *Plant Cell* **16**:5142–5153.

Nakamura Y, Andrés F, Kanehara K, Liu YC, Dörmann, C (2014): Arabidopsis florigen FT binds to diurnally oscillating phospholipids that accelerate flowering. *Nat Commun* **5**: 3553.

Pérez AC, Goossens A. (2013): Jasmonate signalling: a copycat of auxin signalling? *Plant Cell Environ* **36**: 2071–2084.

Pimenta-Lange MJ, Lange T (2006): Gibberellin biosynthesis and the regulation of plant development. *Plant Biol* **8**:281–290.

Quint M, Gray WM (2006): Auxin signaling. *Curr Opin Plant Biol* **9**:448–453.

Rockwell NC, Su YS, Lagarias JC (2006): Phytochrome structure and signaling mechanisms. *Annu Rev Plant Biol* **57**:837–858.

Schaller GE, Bishopp A, Kieber JJ. (2015): The Yin-Yang of hormones: cytokinin and auxin interactions in plant development. *Plant Cell* **27**:44–63.

Shani E, Yanai O, Ori N (2006): The role of hormones in shoot apical meristem function. *Curr Opin Plant Biol* **9**:484–489.

Singh AP, Savaldi-Goldstein S (2015): Growth control: brassinosteroid activity gets context. *J Exp Bot* **66**: 1123–1132.

Traas J, Bohn-Courseau I (2005): Cell proliferation patterns at the shoot apical meristem. *Curr Opin Plant Biol* **8**:587–592.

Twell D (2006): A blossoming romance: gamete interactions in flowering plants. *Nat Cell Biol* **8**:14–16.

Weijers D, Jurgens G (2005): Auxin and embryo axis formation: the ends in sight? *Curr Opin Plant Biol* **8**:32–37.

Willemsen V, Scheres B (2004): Mechanisms of pattern formation in plant embryogenesis. *Annu Rev Genet* **38**:587–614.

Wilson ZA, Yang C (2004): Plant gametogenesis: conservation and contrasts in development. *Reproduction* **128**:483–492.

Yang C, Lu X, Ma B, Chen SY, Zhang JS (2015): Ethylene signaling in rice and *Arabidopsis*: conserved and diverged aspects. *Mol Plant* **8**:495–505.

Zhang Y, Jiao Y, Liu Z, Zhu YX (2015):ROW1 maintains quiescent centre identity by confining WOX5 expression to specific cells. *Nat Commun* **6**:6003.

Zhu JY, Sae-Seaw J, Wang ZY (2013): Brassinosteroid signalling. *Development*, **140**: 1615–1620.

CHAPTER 5

Tissue Culture: The Manipulation of Plant Development

VINITHA CARDOZA

BASF Plant Science LP, Research Triangle Park, North Carolina

5.0. CHAPTER SUMMARY AND OBJECTIVES

5.0.1. Summary

Unique in biology, plant cells are totipotent; whole plants can be regenerated from single nonsexual cells. Tissue culture is a necessary precursor to most plant transformation systems since there must be methods established to manipulate plant tissues and cells in sterile media. Developmental processes can be manipulated in plant tissues by various components included in sterile tissue culture media. These components include sugar, macronutrients, micronutrients, and phytohormones. Developmental processes include callus and somatic embryo formation, shoot formation, and root formation. Tissue culture is not only a necessary enabling technology for transgenic plant production, but it is also used for *in vitro* propagation of valuable plants.

5.0.2. Discussion Questions

1. Differentiate between organogenesis and somatic embryogenesis.
2. Name plant growth hormones used to manipulate tissues *in vitro*.
3. How can tissue culture be used to propagate virus-free plants?
4. What is callus? What are the uses of callus in tissue culture methods?
5. What are protoplasts, and what are their uses?
6. How can haploid plants be produced using tissue culture? Why are haploid cultures useful?

5.1. INTRODUCTION

Plant tissue culture is the *in vitro* (literally "under glass") manipulation of plant cells and tissues, which is a keystone in the foundation of plant biotechnology. It is useful for plant propagation and the study of plant hormones, and is generally required to manipulate and regenerate transgenic plants. Whole plants can be regenerated *in vitro* using tissues, cells, or, sometimes, even a single cell

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

to form whole plants by culturing them on a nutrient medium in a sterile environment. Elite varieties can be clonally propagated, endangered plants can be conserved, virus-free plants can be produced by meristem culture, germplasm can be conserved, and secondary metabolites can be produced. In addition, tissue culture serves as an indispensable tool for transgenic plant production. For nearly any transformation system, an efficient regeneration protocol is imperative. This can be attributed to *totipotency* of plant cells and manipulation of the growth medium and hormones. Plant cells are unique in the sense that every cell has the potential to form whole new plantlike stem cells (stem cell production in mammals is located in time and space, and most mammalian cells cannot be converted to stem cells). However, having an understanding of each plant species and *explant* (donor tissue that is placed in culture) is essential to the development of an efficient regeneration system. The physiological stage of the explant plays a very important role in its response to tissue culture. For example, young explants generally respond better than do older ones.

This chapter examines the history and uses of plant tissue culture and shows how it is integral to plant biotechnology. Furthermore, this chapter presents basic information needed to understand how tissue culture media, including phytohormones, affects plant developmental programs. We will also explore various culture types and regeneration systems. Some people consider tissue culture as more of an art than science since the researcher must develop an eye for differentiating between good and bad (useful and nonuseful) cultures, which has often proved to be the difference between success and failure in plant biotechnology.

5.2. HISTORY OF TISSUE CULTURE

The history of plant tissue culture dates back at least to 1902, when Gottlieb Haberlandt, a German botanist, proposed that single plant cells could be cultured *in vitro*. He tried to culture leaf mesophyll cells but did not have much success. Roger J. Gautheret, a French scientist, had encouraging results with culturing cambial tissues of carrot in 1934. The first plant growth hormone, indole acetic acid (IAA) was discovered in the mid-1930s by F. Kogl and his coworkers. In 1934, Professor Philip White successfully cultured tomato roots. In 1939, Gautheret successfully cultured carrot tissue. Both Gautheret and White were able to maintain the cultures for about 6 years by subculturing them on fresh media. These experiments demonstrated that cultures could not only be initiated but also be maintained over a long period of time. Later in 1955, Carlos Miller and Folke Skoog published their discovery of the hormone kinetin, a cytokinin. Recall from Chapter 4 that cytokinin is an important class of plant growth regulators.

In 1962, Toshio Murashige and Skoog published the composition a plant tissue culture medium known as *MS* (named for the first letters of their last names) *medium*, which now is the most widely used medium for tissue culture. Murashige was a doctoral student in Professor Skoog's lab, and they developed the now-famous MS medium working with tobacco tissue cultures. The formulation of MS medium took place while they were trying to discover new hormones from tobacco leaf extracts, which, when added to tissue cultures, enabled better growth. In a sense, their experiments could be deemed failures since they did not discover a new hormone. Nonetheless, they came up with a seemingly ideal medium for most plant tissue culture work that is used in practically every plant biotechnology laboratory around the world. This major breakthrough in the field of plant tissue culture has enabled nearly all the other breakthroughs cataloged in this book. MS medium seems to be ideal for many cultures since it has all the nutrients that plants require for growth and contains them in the proper relative ratios. The medium has high macronutrients, sufficient micronutrients, and iron in the slowly available chelated form. The success of tobacco culture using MS medium laid the foundation for future tissue culture work, and this has now become the medium of choice for most tissue culture work. MS medium has been improved on in the past 54 years, but the article by Murashige and Skoog (1962) remains one of the most highly cited papers in plant biology.

5.3. MEDIA AND CULTURE CONDITIONS

5.3.1. Basal Media

The success of tissue culture lies in the composition of the growth medium, hormones, and culture conditions such as temperature, pH, light, and humidity. The growth medium is a composition of essential minerals and vitamins that are necessary for a plant's growth and development, including sugar, which the plant needs to thrive—all must be in sterile or axenic conditions. The minerals consist of macronutrients such as nitrogen, potassium, phosphorus, calcium, magnesium, and sulfur, and micronutrients such as iron, manganese, zinc, boron, copper, molybdenum, and cobalt. Iron is seldom added directly to the medium, it is chelated with ethylenediaminetetraacetic acid (EDTA) so that it is more stable in culture and can be absorbed by plants over a wide pH range. Note that EDTA is used in many foods as a preservative. If iron is not chelated with EDTA, it forms a precipitate, especially in alkaline pH. Vitamins are necessary for the healthy growth of plant cultures. The vitamins used are thiamine (vitamin B_1), pyridoxine (B_6), nicotinic acid (niacin), and thiamine. Other vitamins such as biotin, folic acid, ascorbic acid (vitamin C), and vitamin E (tocopherol) are sometimes added to media formulations. Myoinositol, a sugar alcohol, is added to most plant culture media to improve the growth of cultures. In addition, plants require an external carbon source—sugar—since cultures grown *in vitro* rarely photosynthesize sufficiently to support the tissues' carbon needs. Sometimes, cultures are grown in the dark and do not photosynthesize at all. The most commonly used carbon source is sucrose. Other sources used are glucose, maltose, and sorbitol. The pH of the medium is important since it influences the uptake of various components of the medium as well as regulates a wide range of biochemical reactions occurring in plant tissue cultures (Owen et al. 1991). Most media are adjusted to a pH of 5.2–5.8. The acidic pH does not seem to negatively affect plant tissues but delays the growth of many potential contaminants. However, a higher pH is required for certain cultures. Cultures can be grown in either liquid or solidified medium (Fig. 5.1). The medium is most often solidified as it provides a support system for the explants and is easier to handle. *Explant* is the term denoting the starting plant parts used in tissue culture. Solidification is done using agar derived from seaweed or agar substitutes such as Gelrite™ or Phytagel™ commercially available as a variety of gellan gums. These are much clearer than agar. Membrane rafts or filter paper (Fig. 5.1) is also used for support on liquid or semi-solidified medium.



Figure 5.1. Tissue cultures in liquid and solid culture medium. See filter paper bridge in liquid medium.

In addition to MS, there are a plethora of media formulations that are available for plant tissue culture (White 1963; Linsmaier and Skoog 1965; Gamborg et al. 1968; Schenk and Hildebrandt 1972; McCown and Lloyd 1981). McCown's woody plant medium (WPM) has been widely used for tree tissue culture. Knudson's medium (Knudson 1946) was developed for orchid tissue culture and is also used for fern tissue culture. With so many choices in media formulations, one might wonder about how to choose a medium to culture a particular plant species of interest. The choice of medium is typically determined empirically for optimal response of the plant species and the research goal. Typically, explant type and the plant taxonomy are good starting points. For example, nearly all tissue cultures of plants in the Solanaceae (the nightshade family) use MS media. Recall that MS media was developed using tobacco, a member of this plant family. Many times a mix-and-match scheme of macro- and micronutrients from one medium and vitamins from another has also been successful. The composition of nutrients varies from medium to medium. For example, MS medium has higher macronutrients than does WPM. Woody plants typically respond better to WPM than MS medium. It is important to select the right medium for culture according to how the plant empirically responds in tissue culture.

5.3.2. Growth Regulators

The basal medium (e.g., MS) should keep plant tissues alive and thriving. Plant growth regulators (phytohormones or, simply, hormones) are needed to manipulate the developmental program of plant tissues—say, to make callus tissue proliferate, or produce roots from shoots. Growth regulators are the items most often manipulated as experimental factors to enhance tissue culture conditions. The most important growth regulators for tissue culture are auxins, cytokinins, and gibberellins. Both natural and synthetic auxins and cytokinins are used in tissue culture. Auxins promote cell growth and root growth. The most commonly used auxins are indoleacetic acid (IAA), indolebutyric acid (IBA), naphthaleneacetic acid (NAA), and 2,4-dichlorophenoxyacetic acid (2,4-D). An auxin-like compound thidiazuron (TDZ) has increased success rate of plant regeneration in many species. Cytokinins promote cell division and shoot growth. The most commonly used cytokinins are benzylaminopurine (BAP), zeatin, and kinetin. Zeatin is commercially available as zeatin, zeatin riboside, and trans-zeatin. In addition to auxins and cytokinins, other hormones such as abscisic acid (ABA) (Augustine and D'Souza 1997; Cardoza and D'Souza 2002), jasmonic acid (JA) (Blázquez et al. 2004), and salicylic acid (SA) (Hao et al. 2006; Galal 2012) have also been used in plant cell culture. Adjuvants (additional components that enhance growth) that have known to have a positive effect on morphogenesis, such as polyamines including spermidine, spermine, and putrescene, have been used in tissue culture (El Hadrami and D'Auzac 1992; Potdar et al. 1999; Cardoza and D'Souza 2002). By manipulating the amount and combination of growth hormones, regeneration of whole plants from small tissues is possible (Fig. 5.2). A more recently used plant growth regulator brassinolide (BR), a subclass of brassinosteroids, has shown to be effective in plant tissue culture. Brassinosteroids mostly regulate cell elongation. BRs have been shown to promote adventitious shoot regeneration from cauliflower hypocotyls (Sasaki 2002) and improved embryogenic tissue initiation in conifers and rice (Pullman et al. 2003). Over 60 kinds of brassinosteroids have been found to date. Of these, epibrassinolide and homobrassinolide have been mostly used in plant tissue cultures. The structures of various plant growth regulators is shown in Figures 5.3, 5.4, and 5.5 (http://www.sigmaaldrich.com/life-science/molecular-biology/plant-biotechnology/tissue-culture-protocols/growth-regulators.html).

Another critical aspect in plant tissue cultures is the management of the gaseous plant hormone ethylene. When plants are grown *in vitro* in closed culture vessels, there is a buildup of ethylene, which is typically detrimental to the cultures. The addition of ethylene biosynthetic inhibitors such as silver nitrate (Giridhar et al. 2001), aminoethoxyvinylglycine (AVG), and silver thiosulphate (Reis et al. 2003) have been shown to increase the formation of shoots.

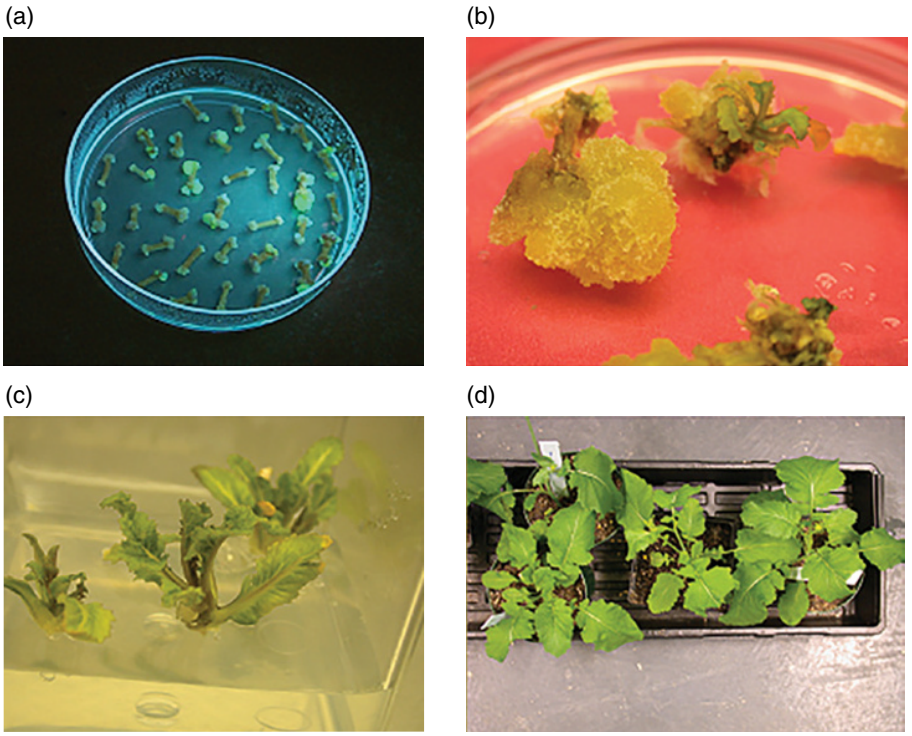


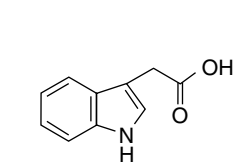
Figure 5.2. *Brassica juncea* plants produced from hypocotyls explants. Shoots are produced when a combination of auxin and cytokinin is used, which is a critical step. The key tissue culture stages for this system is (a) callus from hypocotyl explants; (b) shoots from callus; (c) elongating shoots; and (d) whole plantlets that have been transferred to pots. (*See insert for color representation of the figure.*)

Tissues are transferred to fresh media periodically, say, weekly to monthly, depending on the species and experiment. Without subculturing, tissues will deplete the media of its components and, as the tissues grow they can crowd each other, competing for decreasing resources.

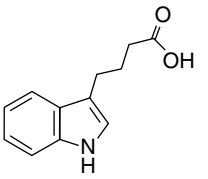
5.4. STERILE TECHNIQUE

5.4.1. Clean Equipment

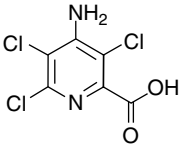
Successful tissue culture requires the maintenance of a sterile environment. All tissue culture work is done in a laminar flow hood. The laminar flow hood filters air with a dust filter and a high-efficiency particulate air (HEPA) filter (Fig. 5.6). It is important to keep the hood clean, which can be done by wiping it with 70% alcohol. The instruments used should also be dipped in 70% ethanol and sterilized using flame or glass beads. Hands should be disinfected with ethanol before handling cultures in order to avoid contamination. It is imperative to maintain axenic conditions throughout the life of cultures: from explant to the production of whole plants. Entire experiments have been lost because of an episode of fungal or bacterial contamination at any stage of culture (see Fig. 5.7). Especially problematic are fungal contaminants that are propagated by spores that might blow into a hood from an environmental source. Therefore, it is important to work away from the unsterile edge of a laminar flow hood. Culture rooms or chambers must be maintained as clean as possible to control any airborne contaminants.



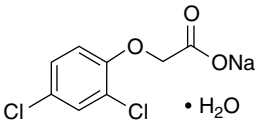
IAA (3-Indoleacetic acid)



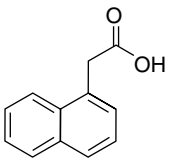
IBA (Indole-3-butyric acid)



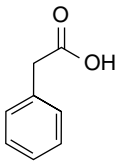
Picloram (4-Amino-3,5, 6-trichloropicolinic acid)



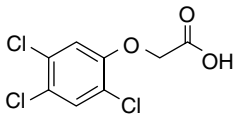
2,4-dichlorophenoxyacetic acid



NAA (1-Naphthaleneacetic acid)



PAA (Phenylacetic acid)

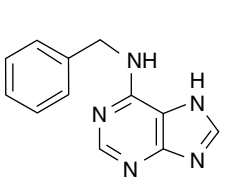


2,4,5-Trichlorophenoxyacetic acid

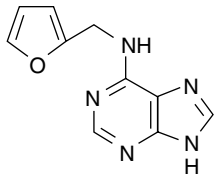
Figure 5.3. Structures of natural and synthetic auxins used in tissue culture.

5.4.2. Surface Sterilization of Explants

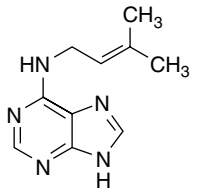
Plant tissues inherently have various bacteria and fungi on their surfaces. It is important that the explant be devoid of any surface contaminants prior to tissue culture since contaminants can grow in the culture medium, rendering the culture nonsterile. In addition, they compete with the plant tissue for nutrition, thus depriving the plant tissue of nutrients. Bacteria and especially fungi can rapidly overtake plant tissues and kill them (Fig. 5.7). The surface sterilants chosen for an experiment typically depend on the type of explant and also plant species. Explants are commonly surface-sterilized using sodium hypochlorite (household bleach), ethanol, and fungicides when using field-grown tissues. The time of sterilization is dependent on the type of tissue; for example, leaf tissue will require a shorter sterilization time than will seeds with a tough seed coat. Wetting agents such as the detergent "Tween" added to the sterilant can improve surface contact with the tissue. Although surface contamination can be eliminated by sterilization, it is very difficult to remove contaminants that are present inside the explant that may show up at a later stage in culture. This internal contamination can be controlled to a certain extent by frequent transfer to fresh medium or by the use of a low concentration of antibiotics in the medium. Overexposing tissues to decontaminating chemicals can also kill tissues, so there is a balancing act between sterilizing explants and killing the explants themselves.



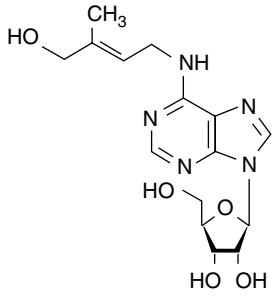
BAP (6-Benzylaminopurine)



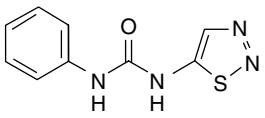
Kinetin (6-Furfurylaminopurine)



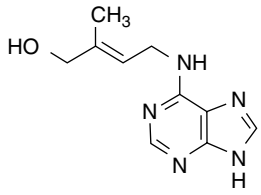
2-iP (N^6-(2-Isopentenyl)adenine)
(6-(Y,Y-Dimethylallylamino)purine)



Zeatin riboside



TDZ (Thidiazuron)



Zeatin (6-(4-Hydroxy-3-methylbut-2-enylamino)
purine

Figure 5.4. Structures of natural and synthetic cytokinins used in tissue culture.

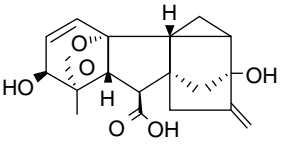
5.5. CULTURE CONDITIONS AND VESSELS

Cultures are typically grown in walk-in growth rooms (Fig. 5.8) or growth chambers. Humidity, light, and temperature have to be controlled for proper growth of cultures. A 16-h light photoperiod is optimal for tissue cultures, and a temperature of 22–25°C is used in most laboratories. A light intensity of 25–50 $\mu mol\, m^{-2}\, s^{-1}$ is typical for tissue cultures and is supplied by cool white fluorescent lamps. A target relative humidity of 50–60% is typically maintained in the growth chambers. Depending on the plant species and the tissue culture types, some cultures are also incubated in the dark. Cultures can be grown in various kinds of vessels such as Petri plates, test tubes, "Magenta" GA7 boxes, bottles, and flasks (Fig. 5.9).

5.6. CULTURE TYPES AND THEIR USES

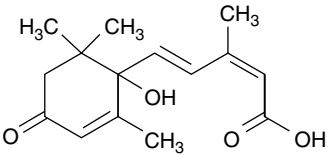
5.6.1. Callus and Somatic Embryo Culture

Callus is an unorganized mass of cells that develops when cells are wounded and is very useful for many *in vitro* cultures. Callus is developed when the explant is cultured on media conducive to undifferentiated cell production—usually the absence of *organogenesis* (organ production) can lead

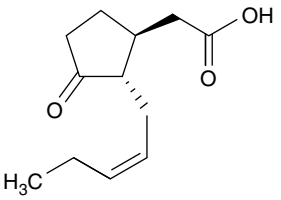


KH

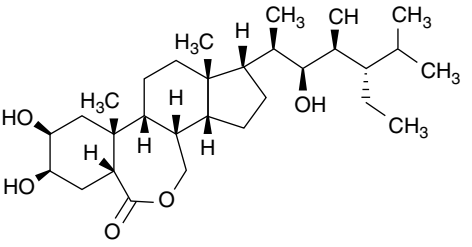
Gibberellic acid (GA)



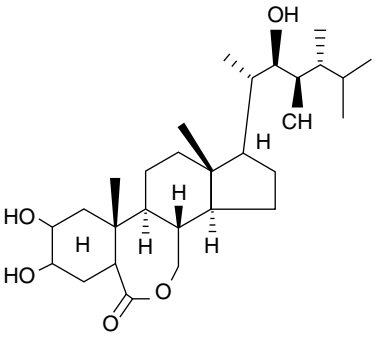
Abscisic acid (ABA)



Jasmonic acid (JA)



22(S),23(S)-Homobrassinolide



Epibrassinolide

Figure 5.5. Structures of other plant growth regulators used in plant tissue culture.



Figure 5.6. Researcher working with tissue cultures in a laminar flow hood. It is important to maintain cultures close to the HEPA filter (left) and away from the edge of the table.

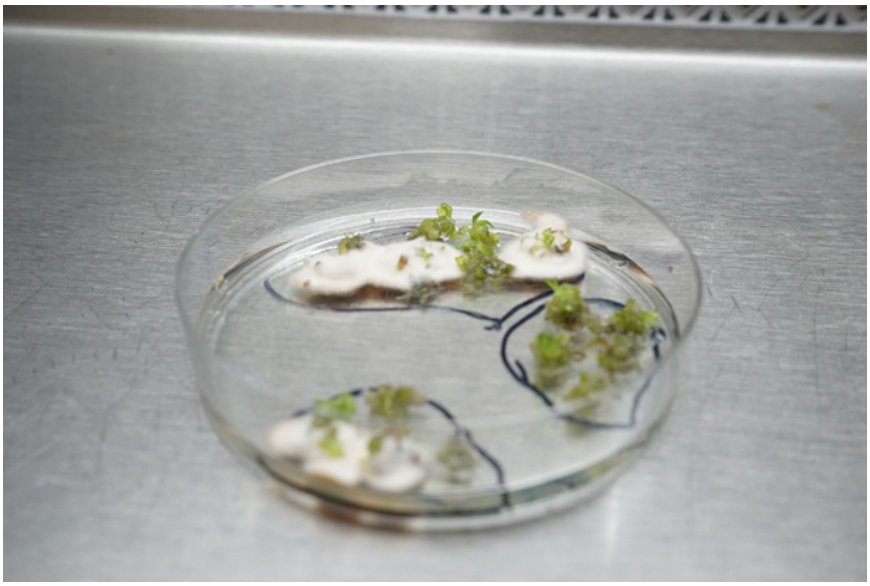


Figure 5.7. Bacteria and fungi contamination of tissue cultures destroy the cultures. In two of the three parts of the Petri dish, the plant cultures have been overgrown by fungal contaminants.



Figure 5.8. A walk-in tissue culture growth room with supplementary cooling and shelves with cool white fluorescent lamps.

to callus proliferation. Stated another way, callus production often leads to organogenesis, but once callus begins to form organs, callus production is halted. Auxins and cytokinins both aid in the formation of most callus cells. Callus can be continuously proliferated using plant growth hormones or then directed to form organs (Fig. 5.10) or somatic embryos (Fig. 5.11)—these systems will be explored in more detail later in this chapter. Callus cultures can be transferred to a new medium for organogenesis or embryogenesis or maintained as callus in culture. Although callus has been induced for various reasons, one important application of callus is to induce *somaclonal* variation through which desired mutants can be selected.

Plant cells can undergo varying degrees of cytological and genetic changes during *in vitro* growth. Some of the changes are derived from preexisting aberrant cells in the explants used for culture.

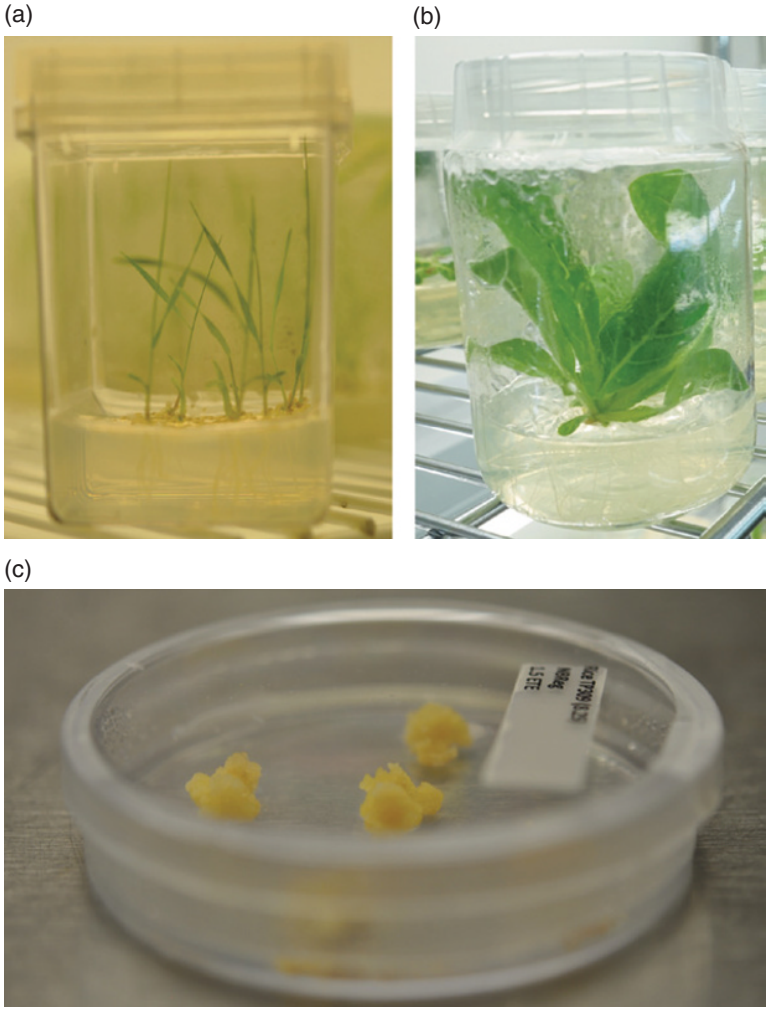


Figure 5.9. Tissue cultures can be grown in various kinds of vessels. Shown here are (a) a Magenta GA7 box, (b) a baby food jar, and (c) a Petri plate—each containing solidified media.

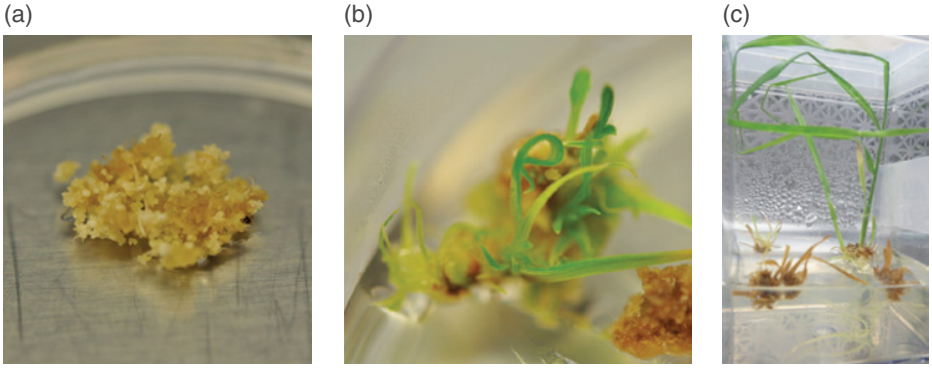


Figure 5.10. Organogenic callus-based system showing (a) callus, (b) callus with early shoot development, and (c) callus with more developed shoots.

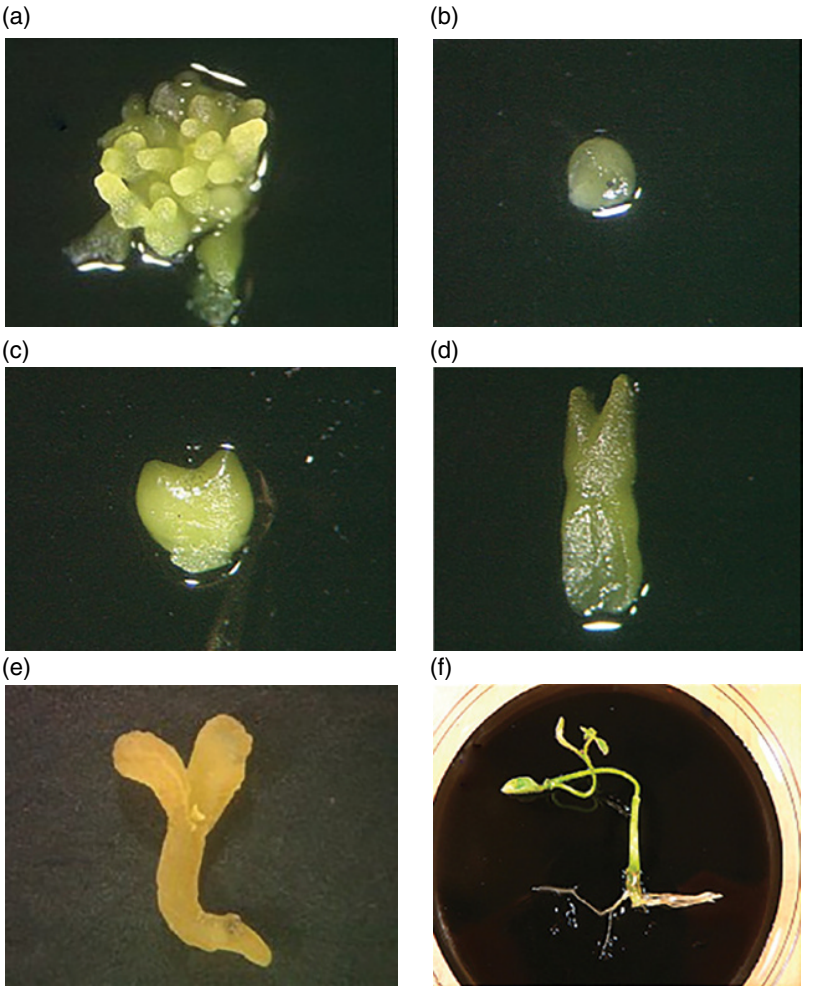


Figure 5.11. Somatic embryogenesis system showing the sequential developmental stages: (a) a cluster of globular somatic embryos, (b) a globular embryo, (c) a heart-shaped embryo, (d) a torpedo-shaped embryo, (e) a mature embryo with cotyledons, and (f) a plantlet from a germinated embryo.

Other changes represent transient physiological and developmental disturbances caused by culture environments. Still other changes can be the result of epigenetic changes, which can be relatively stable but are not transmitted to the progeny. Some variation can be the result of specific genetic changes or mutations that can be transmitted to the progeny. Such genetically controlled variability is known as *somaclonal variation*. Somaclonal variation serves as both a boon and a bane in tissue culture. It may hamper clonal propagation, but at the same time generate desirable somaclonal variants that can be selected for the development of novel cell lines. Induced somaclonal genetic variability of callus can give rise to genetically variable plantlets regenerated from callus; these changes might be utilized in plant breeding. Salt-tolerant (Ochatt et al. 1999), heavy-metal-tolerant (Chakravarty and Srivastava 1997), disease-resistant (Jones 1990), and herbicide-resistant (Smith and Chaleff 1990) cell lines have been selected via somaclonal mutations using callus tissue.

5.6.2. Cell Suspension Cultures

Loose, friable callus, somatic embryos, or various other tissues can be fragmented into small pieces and grown in a liquid medium to form cell suspension cultures. Cell suspensions can be maintained as batch cultures grown in flasks for long periods of time under constant shaking. Somatic embryos

have been initiated from cell suspension cultures (Augustine and D'Souza 1997). The homogeneity and the fast rate of cell growth make them an attractive source to study various cellular and molecular processes. It is possible to regenerate plants from cell cultures provided the regeneration capability is maintained by subculturing the cultures periodically. Subculturing the cell cultures regularly also allows for maintenance of these cultures for long periods of time. Cell cultures can be cryopreserved for use at a later point of time. Cotton plants regenerated from cryopreserved cell cultures have exhibited normal morphology and growth characteristics (Rajasekaran 1996). Cell cultures however vary in morphology. In switch grass, the cell cultures showed three distinct morphologies: sandy, fine milky, and ultrafine (Mazarei et al. 2011). See Fig. 5.12. Interestingly, the source of these cultures was the same and the only difference was the duration in which nutrients were provided to the cell cultures. Cell cultures can be used to isolate protoplasts (Mazarei et al. 2011), which are amenable to genetic manipulation and regeneration into plants. Cell cultures have also been employed for the production of valuable secondary metabolites and heterologous proteins. One "grail" of suspension cultures is the development of single-cell system that can be engineerable, cryopreservable, and regenerable. These do not exist at this time, but a worthy goal.

One use for plant cell cultures is the controlled production of secondary metabolites and recombinant proteins. *Secondary metabolites* are chemical compounds that are not required by the plant for normal growth and development but are produced in the plant as "byproducts" of cell metabolism. That is not to say that secondary metabolites serve no function to the plant; many do. Some are used for defense mechanism or for reproductive purposes such as color or smell. Some important secondary metabolites present in plants are flavonoids, alkaloids, steroids, tannins, and terpenes. Secondary metabolites have been produced using cell cultures in many plant species and have been reviewed by Rao and Ravishankar (2002). The process can be scaled up and automated using bioreactors for commercial production. Many strategies such as biotransformation, cell permeabilization, elicitation, and immobilization have been used to make cell suspension cultures more efficient in the production of secondary metabolites. Secondary metabolite production can be increased by metabolic engineering, in which enzymes in the pathway of a specific compound can be overexpressed together, thereby increasing the production of a specific compound.

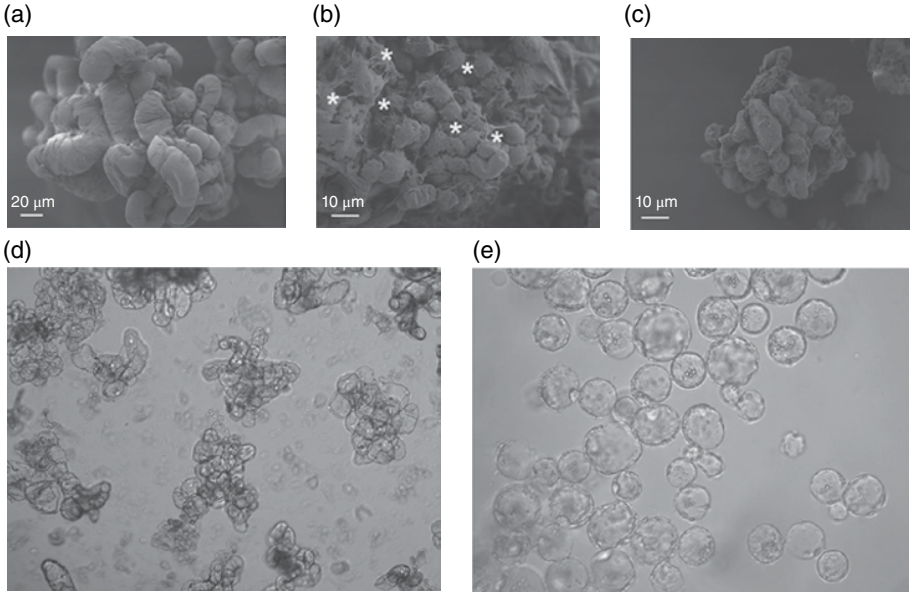


Figure 5.12. Switchgrass cell suspension types and protoplasts isolation. Scanning electron micrographs of switchgrass cell suspension types include (a) sandy, (b) fine milky, and (c) ultrafine. The asterisks indicate extracellular matrix-like layer on the surface of the fine milky cells. (d) Switchgrass cell suspension cultures of the fine milky type before digestion (10×) and (e) protoplasts isolated after digestion (20×). (Reproduced with permission from Mazarei et al. (2011).)

Transgenic plant cell cultures are gaining popularity in the large-scale production of recombinant proteins, thus making them integral parts of molecular farming. What makes molecular farming economically attractive is that production costs can potentially be much lower than those of traditional pharmaceutical production. Plant cell cultures are also advantageous for molecular farming because of high level of containment that they offer relative to whole, field-grown plants and the possibility of commercially producing recombinant proteins. Tobacco suspension culture is the most popular system so far; however, pharmaceutical proteins have been produced in soybean (Smith et al. 2002), tomato (Kwon et al. 2003), and rice (Shin et al. 2003) cells. So far, more than 20 pharmaceutical compounds have been produced in cell suspension cultures, which include antibodies, interleukins, erythropoietin, human granulocyte-macrophage colony-stimulating factor (hGM-CSF), and hepatitis B antigen (Shadwick and Doran 2005).

5.6.3. Anther/Microspore Culture

The culture of anthers or isolated micropsores to produce haploid plants is known as *anther culture* or *microspore culture*. Microspore culture has developed into a powerful tool in plant breeding. Embryos can be produced via a callus phase or be a direct recapitulation of the developmental stages characteristic of zygotic embryos (Palmer and Keller 1997) (Fig. 5.13). It has been known that late uninucleate to early binucleate microspores are the best explants for embryogenesis. In this case, the somatic embryos (explained in Section 5.7.2) develop into haploid plants. Doubled haploids can then be produced by chromosome-doubling techniques. Thus microspore culture enables the production of homozygous (at every locus) plants in a relatively short period as compared to conventional breeding techniques. These homozygous plants are useful tools in plant breeding and genetic studies. In addition, haploid embryos are used in mutant isolation, gene transfer, studies of storage product biochemistry, and physiological aspects of embryo maturation (Palmer and Keller 1997).

5.6.4. Protoplast Culture

Protoplasts contain all the components of a plant cell except for the cell wall (Fig. 5.14). In a few cases using protoplasts, it is possible to regenerate whole plants from single cells and also develop somatic hybrids as described later. Protoplasts offer the possibility of efficient and direct gene transfer to plant cells. DNA uptake has been found to be easier in protoplasts than into intact



Figure 5.13. Somatic embryos regenerated from an anther in culture.

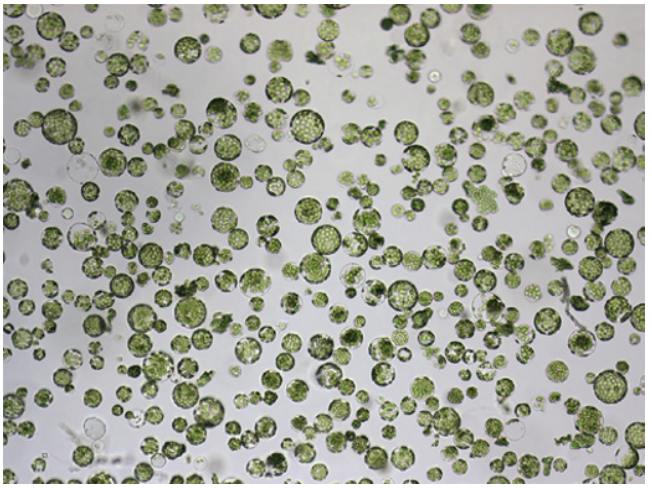


Figure 5.14. Protoplasts derived from the leaves of *Arabidopsis*.

plant cells. Although protoplasts seem to be a very attractive means for plant regeneration and gene transfer, they are very vulnerable to handling. One has to be very careful when manipulating protoplasts. They have to be cultured on a medium with a high osmoticum such as sugars or sugar alcohols; otherwise, the protoplasts will burst open. Protoplasts can be relatively fragile. Furthermore, the removal of cell walls can also disrupt and/or limit developmental programs. Taken together, plant regeneration from protoplasts has proven to be challenging in most cases.

Cell walls are removed from explant tissue mechanically or enzymatically to produce protoplasts. Enzymatic digestion is most common. Enzymatic cell wall degradation was pioneered by Cocking (1960). Since then, protoplast production has been applied to various crop and tree species. Plant cell walls consist of cellulose, hemicellulose, and pectin, with lesser amounts of protein and lipid (Dodds and Roberts 1995). Hence, a mixture of enzymes is necessary for degrading the cell wall. The enzymes that are commonly used are cellulase and pectinase. Following enzyme treatment, protoplasts are purified from cellular debris by filtering using a mesh and then flotation on either sucrose or Ficoll. They are cultured in a high-osmoticum medium to avoid bursting. Protoplasts are cultured either on liquid or solid medium. Protoplasts embedded in an alginate matrix or agarose beads and then cultured on solid or liquid medium have better success rates of regeneration. The alginate or the agarose provides cellular protection against mechanical stress and gradients in environmental conditions during the critical first few days of protoplast culture. Generally, protoplasts first form cell walls and then microcalli, which can be directed for regeneration either toward somatic embryogenesis or toward organogenesis by providing the necessary nutrients.

5.6.5. Somatic Hybridization

Protoplast fusion and somatic hybridization techniques provide the opportunity for bypassing reproductive isolation barriers, thus facilitating gene flow between species. Fusion of protoplasts is accomplished by the use of PEG (poly(ethylene glycol)). Protoplast fusion has helped in the development of somatic hybrids or *cybrids* (cytoplasmic hybrids).

5.6.6. Embryo Culture

Embryo culture is a technique in which isolated embryos from immature ovules or seeds are cultured *in vitro*. This technique has been employed as a useful tool for direct regeneration in species where seeds are dormant, recalcitrant, or abort at early stages of development. Embryo culture also finds

use in the production of interspecific hybrids between inviable crosses, whose seeds are traditionally condemned and discarded because of their inability to germinate. In plant breeding programs, embryo culture goes hand in hand with *in vitro* control of pollination and fertilization to ensure hybrid production. Besides this, immature embryos can be used to produce embryogenic callus and somatic embryos (Ainsley and Aryan 1998) or direct somatic embryos (Cardoza and D'Souza 2000).

5.6.7. Meristem Culture

In addition to being used as a tool for plant propagation, tissue culture is a tool for the production of pathogen-free plants using apical meristem tips. This technique is referred to as *meristem culture, meristem tip culture*, or *shoot tip culture*, depending on the actual explant that is used. Although it is possible to produce bacterium or fungus-free plants, this method has more commonly been used in the elimination of viruses in many species (Kartha and Gamborg 1975; Brown et al. 1988; Ayabe and Sumi 2001). Apical meristems in plants are suitable explants for the production of virus-free plants since the infected plant's meristems typically harbor titers that are either nearly or totally virus-free. Meristem culture in combination with thermotherapy has resulted in successful production of virus-free plants when meristem culture alone is not successful (Kartha 1986; Manganaris et al. 2003; Wang et al. 2006).

5.7. REGENERATION METHODS OF PLANTS IN CULTURE

In plant biotechnology, tissue culture is most important for the regeneration of transgenic plants from single, transformed cells. It is safe to say that without tissue culture there would be no transgenic plants (although this situation is slowly changing—nonetheless tissue culture is required to regenerate intact plants in most species).

5.7.1. Organogenesis

Organogenesis is the formation of organs: either shoot or root. Organogenesis *in vitro* depends on the balance of auxin and cytokinin and the ability of tissue to respond to phytohormones during culture. Organogenesis takes place in three phases. In the first phase, the cells become competent; next, they dedifferentiate. In the third phase, the tissues redifferentiate and morphogenesis proceeds according to the developmental program, which ultimately produce whole intact plants (Sugiyama 1999). *In vitro* organogenesis can be of two types: direct and indirect.

5.7.1.1. Indirect Organogenesis. The formation of organs indirectly via a callus phase is termed *indirect organogenesis*. Induction of plants using this technique does not ensure clonal fidelity, but it could be an ideal system for selecting somaclonal variants of desired characters and also for mass multiplication. Induction of plants via a callus phase has been used for the production of transgenic plants in which (a) the callus is transformed and plants regenerated or (b) the initial explant is transformed and callus and then shoots are developed from the explant.

5.7.1.2. Direct Organogenesis. The production of direct buds or shoots from a tissue with no intervening callus stage is termed *direct organogenesis* (Fig. 5.15). Plants have been propagated by direct organogenesis for improved multiplication rates, production of transgenic plants, and—most importantly—for clonal propagation. Typically, indirect organogenesis is more important for transgenic plant production.

Axillary bud induction/multiple-bud initiation is a set of techniques that represent the most common means of micropropagation since it ensures the production of uniform planting material without genetic variation. Axillary shoots are formed directly from preformed meristems at nodes (Fig. 5.16), and the chance of the organized shoot meristem undergoing mutation is relatively low. This technique is often referred to as multiple-bud induction. Many economically important plants



Figure 5.15. Direct organogenesis that includes shoots and roots from a leaf explant of *Curculigo orchioides.*

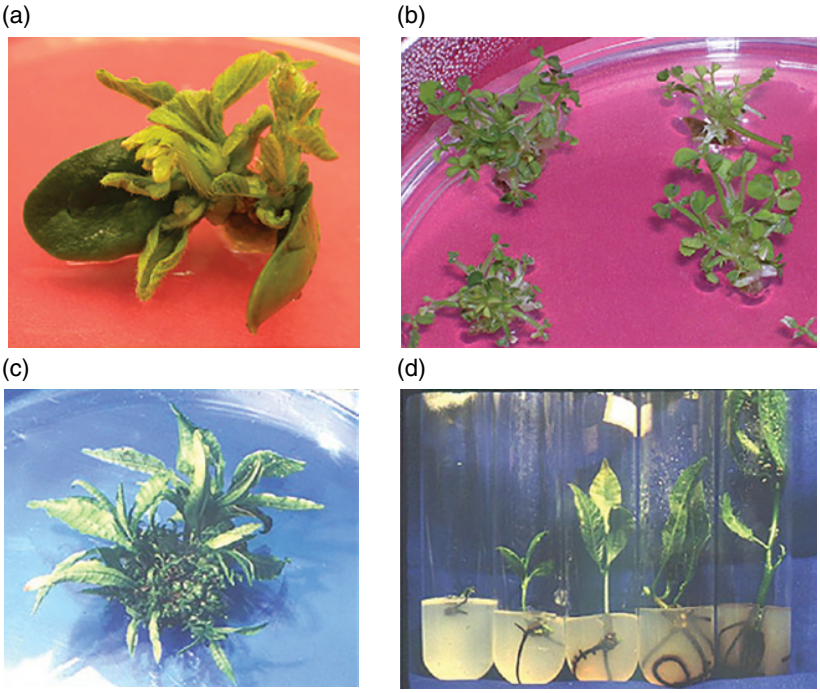


Figure 5.16. Several examples of direct organogenesis in various plant species: (a) multiple bud initiation from cotyledonary nodes of soybean, (b) shoot formation from multiple buds in *Medicago truncatula*, which is a relative of alfalfa, (c) shoot formation from multiple buds of cashew, and (d) the developments of roots and elongating shoots in cashew. (*See insert for color representation of the figure.*)

such as orchids have been propagated using this method. Multiple-bud initiation has been successful in crop plants but in only a few tree species such as *Millingtonia hortensis* (Hegde and D'Souza 1995) and *Fagus sylvatica* (Chalupa 1996). Multiple-bud initiation still remains a challenge in many tree species since many tree species are recalcitrant to tissue culture.

5.7.2. Somatic Embryogenesis

Somatic embryogenesis is a nonsexual developmental process that produces a bipolar embryo with a closed vascular system from somatic tissues of a plant (Fig. 5.11). Somatic embryogenesis has become one of the most powerful techniques in plant tissue culture for mass clonal propagation. Somatic embryogenesis may occur directly or via a callus phase. Direct somatic embryogenesis is preferred for clonal propagation as there is less chance of introducing variation via somaclonal mutation. Indirect somatic embryogenesis is sometimes used in the selection of desired somaclonal variants and for the production of transgenic plants. Large-scale production of somatic embryos using bioreactors and synthetic seeds from somatic embryos has been successful. Somatic embryos can be cryopreserved as synthetic seeds and germinated whenever necessary. One advantage of somatic embryogenesis is that somatic embryos can be directly germinated into viable plants without organogenesis; thus, it mimics the natural germination process.

5.7.3. Synthetic Seeds

Encapsulated somatic embryos are known as *synthetic seeds*. Somatic embryos are typically encapsulated in an alginate matrix, which serves as an artificial seed coat. The encapsulated somatic embryos can be germinated *ex vitro* ("out of glass") or *in vitro* to form plantlets. Synthetic seeds have multiple advantages—they are easy to handle, they can potentially be stored for a long time, and there is potential for scale-up and low cost of production. Another advantage is the prospects for automation of the whole production process because the commercial application of somatic embryogenesis requires high-volume production. Synthetic seeds can be stored at 4°C for shorter periods or cryopreserved in liquid nitrogen for long-term storage (Fang et al. 2004). Production of synthetic seeds and germination of these seeds to plantlets has been accomplished in sandal wood, coffee, bamboo, and many other plant species.

5.8. ROOTING OF SHOOTS

Efficient rooting of *in vitro*-grown shoots is a prerequisite for the success of micropropagation. The success of acclimatization of a plantlet greatly depends on root system production. Rooting of trees and woody species is difficult compared with that of herbaceous species. Rooting of shoots is achieved *in vitro* or *ex vitro*. *Ex vitro* rooting reduces the cost of production significantly. *Ex vitro* rooting is carried out by pretreating shoots with phenols or auxins and then directly planting them in soil under high-humidity conditions. Using this method, acclimation of the rooted shoots can be carried out simultaneously. *In vitro* rooting consists of rooting the plants in axenic conditions. Despite relatively high costs, *in vitro* rooting is still a very common practice in many plant species because of its several advantages. Tissue culture facilitates administration of auxins and other compounds, avoids microbial degradation of applied compounds, allows the addition of inorganic nutrients and carbohydrates, and enables experiments with small, simple explants. Several factors are known to affect rooting. The most important factor is the action of endogenous and exogenous auxins. In many cases, a pulse treatment with auxins for a short period has also been sufficient for root induction.

Phenolic compounds are known to have a stimulatory effect on rooting. Among the phenolic compounds, phloroglucinol, known as a root promoter, has a positive effect on rooting (Hegedus and Phan 1987). Catechol, a strong reducing agent, has been reported to regulate IAA oxidation and thus affects rooting in plant tissue culture (Hackett 1970).

5.9. ACCLIMATION

Once plants are generated by tissue culture, they have to be transferred to the greenhouse or field. This requires that the plants be hardened-off before transfer to the field. During this acclimation process, plants are first transferred to a growth chamber or greenhouse and covered by domes to minimize the loss of water. Tissue culture conditions are at approximately 100% humidity, whereas the relative humidity outside the vessels is typically much lower. In addition, the plants must be "weaned" off the rich media so they can grow as normal plants in soil. Once plants are acclimated under greenhouse conditions, they are ready for transfer to the field. Acclimation is a very important step in tissue culture, because plants might die if they are not properly hardened-off.

5.10. PROBLEMS THAT CAN OCCUR IN TISSUE CULTURE

5.10.1. Culture Contamination

A serious and most frequent problem encountered by plant tissue culturists is the presence of covert microbial contaminants in the cultures. This is the main cause of losses in commercial and scientific plant tissue culture laboratories. Contamination can result in the death of cultures, growth retardation, necrosis, and altered morphogenetic potential such as reduced rate of multiplication and rooting. Both epiphytic and endophytic organisms can cause severe losses to micropropagated plants at each stage of growth. Contamination is not always seen in the culture establishment stage; some internal contaminants become evident at later subcultures and are difficult to eliminate as a result of poor handling of cultures. Proper aseptic techniques can reduce contamination as a result of improper handling of cultures. Controlling contamination by proper sterilization techniques is discussed in Section 5.4.2.

5.10.2. Hyperhydricity

Hyperhydricity, formerly known as vitrification, is a common problem encountered in tissue culture. Hyperhydricity is the excessive hydration of tissue cultured plants that results in the reduced mechanical strength of these plants and plant regeneration. The success rate of acclimating a hyperhydrated plant is low compared to a healthy plant. These plants have a translucent appearance with thick, brittle leaves. Microscopic observations have revealed that leaves of hyperhydric plants have a poorly developed epicuticular wax layer, a reduced number of palisade cells, and large intercellular spaces in the mesophyll (Olmos and Hellin 1998; Picoli et al. 2001; Jausoro et al. 2010). Flooding of the apoplast seems to be a key factor in the development of hyperhydricity (Dries et al. 2013). The causative agents of hyperhydricity may be high relative humidity in the culture vessels, gas accumulation such as ethylene in the vessels, concentration and type of gelling agent in the media, and other factors that trigger oxidative stresses such as high salt concentration and low light intensity. Hyperhydricity can be controlled by adjusting the relative humidity in the culture vessels. An environment that allows proper exchange of gases and water vapor are helpful in reducing hyperhydricity. Using higher concentration of a gelling agent also reduces hyperhydricity. The use of $AgNO_3$ an ethylene inhibitor has been reported to reduce hyperhydricity in sunflower (Mayor et al. 2003) and potato (Sandra and Maira 2013).

5.10.3. Browning of Explants

A major constraint in tissue culture of woody plants is the browning of explants, resulting in tissue death. In this case, cut surfaces start to discolor soon after excision and continue exuding phenolics even after culturing. Browning is the result of a mixture of complex phenolic exudates often found

at high levels in woody plants. These phenols become toxic by being reversibly hydrogen-bonded to proteins and oxidated to form highly active quinones which then become cyclic or polymerized and oxidize proteins to form increasingly melanic compounds called "polyphenols" (Harms et al. 1983).

Pretreatment with antioxidants has been used to control browning in *in vitro* cultures. The most commonly used antioxidants are ascorbic acid and citric acid. Antioxidants are used either added to media (Ko et al. 2009; Ndakidemi et al. 2014) or as explant pretreatment prior to culturing. Chilled antioxidant treatment of explants also helps in reducing browning of tissues. Addition of certain additives such calcium pantothenate, poylvinylpyrrolidone (PVP), and activated charcoal have also help in reducing browning of tissue cultures. Activated charcoal acts by adsorbing phenolic exudates. Culture conditions such as incubation in the dark or frequent subculture also assist in reducing browning.

5.11. CONCLUSIONS

Plant tissue culture is an essential tool in plant biotechnology that has enabled mass clonal propagation, production of secondary metabolites, preservation of germplasm, and production of virus-free plants. Moreover, it serves as an indispensable tool for regenerating transgenic plants. All these have been possible by manipulating plant tissues and various kinds of media developed by plant tissue culturists and by the use of plant hormones. It has been one of the very exciting discoveries for plant biologists and will continue to be most useful in the coming years.

ACKNOWLEDGMENTS

The author would like to thank Dr. Leo D'Souza from the Laboratory of Applied Biology, St Aloysius College, Mangalore, India, for suggestions on the manuscript and for kindly providing many of the illustrations for this chapter. Many thanks to Dr. Wayne Parrott and Benjamin Martin at the University of Georgia and Jennifer Hinds at the University of Tennessee for providing images used in this chapter.

LIFE BOX 5.1. GLENN BURTON COLLINS

Glenn Burton Collins, Professor Emeritus, Department of Plant and Soil Sciences, University of Kentucky.



Glenn Collins. Courtesy of Glenn Collins.

My interest in science began during my growing-up years living and working on a farm. There, I developed a fascination with the diversity of the plants and animals in our fields and streams. I became even more fascinated and intrigued with living organisms as I began to take science classes and started to understand how living organisms functioned and adapted to their diverse environments. The defining moment that led me down my specific educational and career pathway was in the summer between my sophomore and junior years of college when I got a job working for a plant breeder.

I worked for the same breeder during the remainder of my baccalaureate degree program, and I subsequently did my master's degree under his direction. I then pursued my Ph.D. degree in Genetics with a minor in Plant Physiology at North Carolina State University. Back in those days, we did not typically take postdoctoral appointments unless we had problems getting an offer of a permanent position.

I have been in an academic appointment at the University of Kentucky since completing my Ph.D. degree in 1966. Training in the field of genetics and plant breeding/cytogenetics was a wonderful platform for being positioned to participate in and contribute to the advancements in plant biotechnology. I headed up a team that developed and released ten new cultivars and eleven germplasm lines during the 14 years that I was in my faculty position as the plant breeder. At the same time, my program made major contributions to crop improvement by developing alternate strategies for crop improvement that included improved plant tissue culture systems; producing haploids and doubled haploids from microspores in cultured anthers; and in generating new interspecific hybrid combinations in plants using *in vitro* embryo rescue and protoplast fusion. I moved into a more basic-science-oriented faculty position in 1980 that was defined as plant somatic cell genetics. This position change was well timed for the vast opportunities that were made available by recombinant DNA tools and genetic engineering approaches for putting foreign genes into plants using *in vitro* cultured explants in an aseptic tissue culture environment. We had already developed efficient totipotent *in vitro* systems for several plant species including for several Trifolium species in the legume family. We generated transgenic soybeans in the late 1980s and to date we have introduced genes for disease resistance, herbicide tolerance, and biochemical trait modifications into a number of plant species.

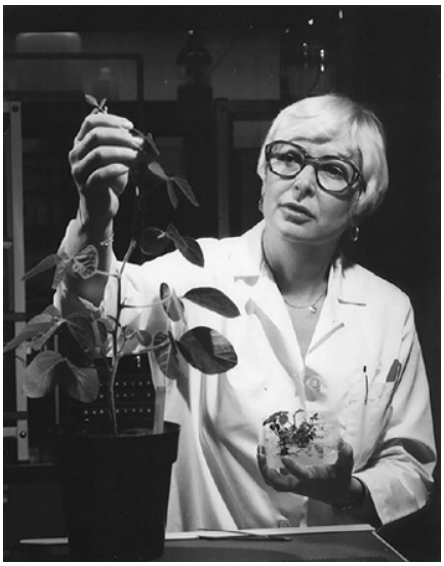
In addition to these cited examples of contributions to the shaping of plant biotechnology, I give a lot of credit to people who have been in or associated with my program and provided major advancements to the field of biotechnology, both while in my laboratory and then in their own career positions. These include 17 doctoral students, 12 M.S. students,25 postdoctoral fellows, 24 visiting scientists, the staff in my laboratory, and my many collaborators.

Another very significant contribution that we have made to the advancement of plant biotechnology has been the training of a very large number of B.S. degree recipients through our interdisciplinary program in agricultural biotechnology, which was initiated in 1988 as a research oriented baccalaureate degree. A majority of these graduates have gone into doctoral and professional degree programs with a substantial number of them in biotechnology careers. Many other graduates have accepted positions in the field of biotechnology with private companies or in university and government laboratories.

I have a difficult time feeling precise and inclusive when I think about trying to predict future advancements in plant biotechnology. The reason is because the advancements are so rapid, numerous, and diverse as we utilize functional genomics and other approaches to identify genes and the traits that they control in plants, that predicting the myriad of applications is mind-boggling. The knowledge-base that will be generated will certainly provide the opportunity to improve crops for their current traditional uses and also to engineer plants for new uses in health, nutrition, energy, and environmental applications.

LIFE BOX 5.2. MARTHA S. WRIGHT

Martha S. Wright, Research Scientist (retired), Syngenta and Monsanto.



Martha Wright with a regenerated soybean plant (1981). Courtesy of Martha Wright.

My love of science emerged in high school when I entered the Kansas City Science Fair in 1956. For my project, I disassembled an animal from each of the phyla and put their skeletons back together for comparison. The project didn't win anything, but my mom was glad that I wasn't boiling lizards on her stove anymore. At Lindenwood College, now Lindenwood University, in St. Charles, Missouri, I originally majored in business because my father said I'd always be able to get a job as a secretary. Remember, this was 1958 and I grew up in Kansas. In my sophomore year, I was fortunate to have an advisor in the business department who noted that I was bored. After a discussion, she urged me to sign up for an advanced biology course. Ultimately, I graduated in 1962 from Lindenwood with a major in biology and minors in chemistry, classics, and business.

During my last 2 years of college, I worked summers in hospital laboratories. My first permanent job after graduating was in the Agricultural Division of Monsanto in St. Louis. I was hired because I had worked with radioactivity while in college. One of my biology professors had worked on the Manhattan Project. My first assignment at Monsanto was to work on an insecticide. For the next 15 years, I worked on a series of projects: some having to do with animals and some with plants. I especially enjoyed my early work with Roundup®. We were able to determine the mode of action through a series of experiments using Lemna as a model system. In the early 1970s, Ernie Jaworski, my supervisor at Monsanto at that time, went to Saskatoon on a sabbatical in Olaf Gamborg's laboratory. When he returned, he handed me some cell cultures and they became mine, and that was the beginning of my true career.

In the early days of field crop cell culture, the "holy grail" was soybean and it was thought to be impossible to regenerate from cell culture. An understanding of the way plant hormones act at different stages was probably the single most important factor to aid soybean regeneration. I was fortunate to work with Michael Carnes as we unraveled the hormone profiles of several field crops, soybean, maize, alfalfa, and so on. By the mid 1980s, we had published three methods for regenerating soybeans from cell culture. Concurrently, molecular biologists were having some success with plant cell transformation.

Eventually, several crops were in the race to the field. I left Monsanto in 1987 and started working at then Ciba-Geigy, now Syngenta, in North Carolina. Here, the emphasis was clearly on corn to support our seed company. Monocots were proving to be especially difficult to transform; and it seemed that if the regeneration system worked, it didn't mesh with the transformation procedure, and vice versa. In 1991, we published the recovery of fertile transformed maize plants using the Biolistics® gun.

During the rest of my career, I either worked directly with, or supervised work

with, soybean, corn, vegetables, cotton, rice, and so on. I learned from every one of the wonderful, intelligent, dedicated people with whom I worked throughout my career, and I am eternally grateful to them for being part of my work and my life. In 2001, I retired from Syngenta. I was happy, healthy, and satisfied with my career. Now I'm giving back in various volunteer capacities.

I feel our work broke the mystique of plant regeneration from cell culture, and ultimately allowed the transformation of recalcitrant crops. Transformation, in most cases, depends upon being able to work at the cellular level, thus without the ability to regenerate from cell culture, recovery of fertile transformants is not possible.

Today, it is critical that we continue the hunt for beneficial crop genes while we look for other crops where we can make a valuable difference. Of equal importance is to educate the public on the safety of enhanced crops. We must convince the nonscientists that we too have children and do not want harm to come to them or anyone, now or in the future. This stance also requires responsibility on our part. Enhanced crops mean more people get to eat and more people are healthy and can devote their energies to improving the world. That's the goal. And that's always been my goal in agricultural research.

LIFE BOX 5.3. VINITHA CARDOZA

Vinitha Cardoza, Senior Scientist, BASF Plant Science, Research Triangle Park, NC.



Vinitha Cardoza. Courtesy of Vinitha Cardoza.

I always wanted to be a scientist from the time I was an elementary student growing up in the coastal Southern town of Mangalore in India. At that point, I was not sure how I would do it, but I knew that was what I wanted to do. Plants always fascinated me and as a little girl, I wondered if I could make the May Flower (*Delonix regia* which is also called "Gulmohar"), which flowers profusely in the month of May in the region where I live, flower all through the year. With my genetics classes in high school, I realized the potential power of genes and developed a passion for genetics.

One college professor, Dr. Leo D'Souza, fed my curiosity for research by teaching me observational skills during his plant anatomy class. He also taught a course in genetics. Learning more about the power of genes intrigued me to a point where I decided that I definitely needed to follow this path. Dr. D'Souza also ran a tissue culture laboratory working on economically important tree species that lived in the area. As an undergraduate student, I frequently visited his laboratory and decided that I would come back to work there. I pursued my masters in Cytogenetics from Bombay University, in the city now known as Mumbai, working on the mutagenesis of chili pepper. I studied the fascinating meiotic and mitotic changes that occur after mutagenesis and related these changes to the plant phenotype. When it was time for my Ph.D., I decided to go back to my alma mater and Dr. Leo's lab where he asked me if I was ready to work on a challenging and difficult topic: tissue culture of cashew, a very recalcitrant species. I accepted the

challenge and was successful in obtaining somatic embryos in cashew and also doing some early work on transformation in cashew with marker genes. During my Ph.D. studies, I also worked on protoplast culture in cashew. I mastered tissue culture here and learned the early skills in plant transformation. As a Ph.D. student, I gave several talks on tissue culture and biotechnology to a wide range of audiences: from farmers to medical doctors. It was then that I developed the competency of presenting my research work. I realized that it is not just important to be a scientist but also to disseminate this information to others in a way that people untrained in science can understand.

After my Ph.D., I was fortunate to find a postdoc position with Dr. Neal Stewart who was then at the University of North Carolina at Greensboro, where I worked on canola and soybean transformation and successfully increased the transformation efficiency of canola. While in his lab, I developed my molecular biology skills and worked on identifying inducible promoters in addition to learning canola and soybean transformation. Dr. Stewart's lab was very diverse with people having different set of skills. Here, I learned a lot from interacting and working with my fellow postdocs and also got the opportunity to mentor graduate students, which I very much enjoyed. My second postdoc at the University of Vermont took me into the field of plant–microbe interactions, where I worked on the molecular and signaling aspects of nodulation in *Medicago truncatula*. While in Vermont, I had the opportunity to teach as well as pursue research. I enjoyed mentoring and teaching experiences. I developed and taught course in plant biotechnology. At Vermont, I had a special postdoc position called the "teaching postdoc" where 25% of my time was devoted to teaching. Such programs help postdocs develop teaching skills with a mentor. My postdoc advisor, Dr. Jeanne Harris, was also my teaching mentor. I enjoyed teaching and fine-tuned my teaching skills there.

When it came time to make a career decision between teaching/academia and industry, I struggled to make the right choice. I enjoyed teaching as much as enjoyed doing research, and I enjoyed the university environment. However, I finally decided to take an industry position as I thought it would be the fastest way to utilize my research skills to deliver a product where I could make an impact to society. In 2006, I joined BASF Plant Science to work in their plant transformation group where I leveraged my transformation and tissue culture skills. However, within a year I had the opportunity to move to the molecular discovery group at the company. Having both transformation and molecular biology skills put me in a very good position to understand the entire chain of starting with DNA and ending with the whole transgenic plant. I enjoy working in industry and contributing to gene discovery at BASF. I feel dreams come true if you follow them and choose the right path. As I mentioned at the beginning, I knew I wanted to be a scientist and I am one now—a dream come true.

REFERENCES

Ainsley PJ, Aryan AP (1998): Efficient plant regeneration system for immature embryos of triticale (*x Triricosecale Wittmack*). *Plant Growth Reg* **24**:23–30.

Augustine AC, D'Souza L (1997): Somatic embryogenesis in *Gentum ula* Brongn. *(Gentum edule)* (Willd) Blume. *Plant Cell Rep* **16**:354–357.

Ayabe M, Sumi S (2001): A novel and efficient tissue culture method—"stem-disc dome culture"—for producing virus-free garlic (*Allium sativum* L.). *Plant Cell Rep* **20**:503–507.

Blázquez S, Piqueras A, Serna MD, Casas JL, Fernández JA (2004): Somatic embryogenesis in saffron: optimization through temporary immersion and polyamine metabolism. *Acta Hort* **650**:269–276.

Brown GR, Kwiatkowski S, Martin MW, Thomas PE (1988): Eradication of PVS from potato clones through excision of meristems from *in vitro* heat-treated shoot tips. *Am Pot J* **65**:633–638.

Cardoza V, D'Souza L (2000): Direct somatic embryogenesis from immature zygotic embryos in cashew *(Anacardium occidentale* L.). *Phytomorphology* **50**(2):201–204.

Cardoza V, D'Souza L (2002): Induction, development and germination of somatic embryos from nucellar tissues of cashew *(Anacardium occidentale* L.). *Sci Hort* **93**:367–372.

Chakravarty B, Srivastava S (1997): Effects of genotype and explant during *in vitro* response to cadmium stress and variation in protein and proline contents in linseed. *Ann Bot* **79**:487–491.

Chalupa V (1996): *Fagus sylvatica* L. (European beech). In *Biotechnology in Agriculture and Forestry, Vol. 35: Trees IV*, Bajaj YPS (ed). Springer-Verlag, Berlin/Heidelberg/New York, pp 138–152.

Cocking EC (1960): A method for the isolation of plant protoplasts and vacuoles. *Nature* **187**: 927–929.

Dodds JH, Roberts LW (1995): Isolation and culture of protoplasts. In *Experiments in Plant Tissue Culture*, 3rd ed, Dodds JH, Roberts LW (eds). Cambridge University Press, Cambridge, UK, pp 167–182.

Dries NV, Gianni S, Czerednik A, Krens FA, Klerk GM (2013): Flooding of the apoplast is a key factor in the development of hyperhydricity. *Exp Bot* **64**: 5221–5230.

El Hadrami I, D'Auzac J (1992): Effects of polyamine biosynthetic inhibitors on somatic embryogenesis and cellular polyamines in *Hevea brasiliensis*. *J Plant Physiol* **140**:33–36.

Fang J, Wetten A, Hadley P (2004): Cryopreservation of cocoa *(Theobroma cacao* L.) somatic embryos for long-term germplasm storage. *Plant Sci* **166**:669–675.

Galal A (2012): Improving the effect of salicylic acid on the multipurpose tree *Ziziphus spina-christi* (L.) wild tissue culture. *Am J Plant Sci* **3**: 947–952.

Gamborg OL, Miller RA, Ojima K (1968): Nutrient requirements of suspension cultures of soybean root cells. *Exp Cell Res* **50**:151–158.

Giridhar P, Reddy BO, Ravishankar GA (2001): Silver nitrate influences *in vitro* shoot multiplication and root formation in *Vanilla planifolia* Andr. *Curr Sci* **81**:1166–1170.

Hackett WP (1970): The influence of auxin, catechol and methanolic tissue extracts on root initiation in aseptically cultured shoot apices of juvenile and adult forms of *Hedera helix*. *J Am Soc Hort Sci* **95**:398–402.

Hao L, Zhou L, Xu X, Xu JC, Xi T (2006):The role of salicylic acid and embryogenic carrot extracts in somatic embryogenesis of naked oat (*Avena nuda*). *Plant Cell Tiss Org Cult* **85**: 109–113.

Harms CT, Baktir I, Oertli JJ (1983): Clonal propagation *in vitro* of red beet (*Beta vulgaris* spp.) by multiple adventitious shoot formation. *Plant Cell Tiss Org Cult* **2**: 93–102.

Hegde S, D'Souza L (1995): *In vitro* propagation of an ornamental tree *Millingtonia hortensis* L.f. *Gartenbauwissenschaft* **60**:258–261.

Hegedus P, Phan CT (1987): Activities of five enzymes of the phenolic metabolism on rooted and acclimatised vitreous plants in relation with phenolic treatments. *Acta Hort* **221**:211–216.

Jausoro V, Llorente BE, Apostolo NM (2010): Structural differences between hyperhydric and normal *in vitro* shoots of *Handroanthus impetiginosus* (Mart. ex DC) Mattos (Bignoniaceae). *Plant Cell Tiss Org Cult* **101**:183–191.

Jones PW (1990): *In vitro* selection for disease resistance. In *Plant Cell Line Selection*, Dix PJ (ed). VCH, Weinheim, pp 113–150.

Kartha KK (1986): Production and indexing of disease free plants. In *Plant Tissue Culture and Its Agricultural Application*, Withers LA, Alderson PG (eds). Butterworths, London, pp 219–238.

Kartha KK, Gamborg OL (1975): Elimination of cassava mosaic disease by meristem culture. *Phytopathology* **65**:826–828.

Knudson L (1946): A new nutrient solution for germination of orchid seed. *Am Orch Soc Bull* **15**:214–217.

Ko WH, Su CC, Chen CL, Chao CP (2009): Control of lethal browning of tissue culture plantlets of Cavendish banana cv. Formosano with ascorbic acid. *Plant Cell Tiss Org Cult* **96**:137–141.

Kwon TH, Kim YS, Lee JH, Yang MS (2003): Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato cell suspension cultures. *Biotechnol Lett* **25**:1571–1574.

Linsmaier EM, Skoog F (1965): Organic growth factor requirements of tobacco tissue cultures. *Physiol Plant* **18**:100–127.

Manganaris GA, Economou AS, Boubourakas IN, Katis NI (2003): Elimination of PPV and PNRSV through thermotherapy and meristem-tip culture in nectarine. *Plant Cell Rep* **22**:195–200.

Mayor ML, Nestares G, Zorzoli R, Picardi LA (2003): Reduction of hyperhydricity in sunflower tissue culture. *Plant Cell Tiss Org Cult* **72**:99–103.

Mazarei M, Al-Ahmad H, Rudis MR, Joyce BL, Stewart CN Jr (2011): Switchgrass (*Panicum virgatum* L.) cell suspension cultures: establishment, characterization, and application. *Plant Sci* **181**:712–715.

McCown BH, Lloyd G (1981): Woody plant medium (WPM)—a mineral nutrient formulation for microculture for woody plant species. *Hort Sci* **16**:453.

Murashige T, Skoog F (1962): A revised medium for rapid growth and bioassay with tobacco tissue cultures. *Physiol Plant* **15**:473–497.

Ndakidemi CF, Mneney E, Ndakidemi PA (2014): Effects of ascorbic acid in controlling lethal browning in *in vitro* culture of *Brahylaena huillensis* using nodal segments. *Am J Plant Sci* **5**:187–191.

Ochatt SJ, Marconi PL, Radice S, Arnozis PA, Caso OH (1999): *In vitro* recurrent selection of potato: production and characterization of salt tolerant cell lines and plants. *Plant Cell Tiss Org Cult* **55**:1–8.

Olmos E, Hellín E (1998): Ultrastructural differences of hyperhydric and normal leaves from regenerated carnation plants. *Sci Hort* **75**:91–101.

Owen HR, Wengaerd D, Miller AR (1991): Culture medium pH is influenced by basal medium, carbohydrate source, gelling agent, activated charcoal, and medium storage method. *Plant Cell Rep* **10**:583–586.

Palmer CE, Keller WA (1997): Pollen embryos. In *Pollen Biotechnology for Crop Production and Improvement*, Shivanna KR, Sawhney VK (eds). Cambridge University Press, Cambridge, UK, pp 392–422.

Picoli EAT, Otoni WC, Figueira ML, Carolino SMB, Almeida RS, Silva EAM, Carvalho CR, Fontes EPB (2001): Hyperhydricity in *in vitro* eggplant regenerated plants: structural characteristics and involvement of BiP (binding protein). *Plant Sci* **160**:857–868.

Potdar UA, Khan BM, Rawal SK (1999): Genotype dependent somatic embryogenesis in sunflower (*Helianthus annuus* L.). In *Plant Tissue Culture and Biotechnology*, Kavi Kishor, PB (ed). University Press, Hyderabad, pp 64–70.

Pullman GS, Zhang, Phan BH (2003): Brassinolide improves embryogenic tissue initiation in conifers and rice. *Plant Cell Rep* **22**:96–104.

Rajasekaran K (1996): Regeneration of plants from cryopreserved embryogenic cell suspension and callus cultures of cotton (*Gossypium hirsutum* L.). *Plant Cell Rep* **11**:859–64.

Rao SR, Ravishankar GA (2002): Plant cell cultures: chemical factories of secondary metabolites. *Biotechnol Adv* **20**:101–153.

Reis LB, Paiva Neto VB, Todedo Picoli EA, Costa MGC, Rego MM, Carvalho CR, Finger FL, Otoni WC (2003): Axillary bud development of passionfruit as affected by ethylene precursor and inhibitors. *In Vitro Cell Dev Biol—Plant* **39**:618–622.

Sandra AT, Maira O (2013): Effect of culture medium consistence and silver nitrate on micropropagation of two potato (*Solanum tuberosum*) cultivars. *Rev Colomb Biotecnol* **15**:55–62.

Sasaki H (2002): Brassinolide promotes adventitious shoot regeneration from cauliflower hypocotyls segments. *Plant Cell Tiss Org Cult* **71**:111–116.

Schenk RU, Hildebrandt AC (1972): Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can J Bot* **50**:199–204.

Shadwick FS, Doran PM (2005): Foreign protein expression using plant cell suspension and hairy root cultures. In *Molecular Farming: Plant-Made Pharmaceuticals and Technical Proteins*, Fischer R, Schillberg S (eds). John Wiley & Sons Inc., New York, pp 13–26.

Shin YJ, Hong SY, Kwon TH, Jang YS, Yang MS (2003): High level of expression of recombinant human granulocyte-macrophage colony stimulating factor in transgenic rice cell suspension cultures. *Biotechnol Bioeng* **82**:778–783.

Smith WA, Chaleff RS (1990): Herbicide resistance. In *Plant Cell Line Selection*, Dix PJ (ed). VCH, Weinheim, pp 151–166.

Smith ML, Mason HS, Shuler ML (2002): Hepatitis B surface antigen (HbsAg) expression in plant cell culture: kinetics of antigen accumulation in batch culture and its intracellular form. *Biotechnol Bioeng* **80**:812–822.

Sugiyama M (1999): Organogenesis *in vitro*. *Curr Opin Plant Biol* **2**:61–64.

Wang L, Wang G, Hong N, Tang R, Deng X, Zhang H (2006): Effect of thermotherapy on elimination of apple stem grooving virus and apple chlorotic leaf spot virus for *In vitro*-cultured pear shoot tips. *Hort Sci* **41**:729–732.

White PR (1963): *The Cultivation of Animal and Plant Cells*, 2nd ed. Ronald Press Company, New York.

CHAPTER 6

Molecular Genetics of Gene Expression

MARIA GALLO

Molecular Biosciences and Bioengineering Department, University of Hawaii at Mānoa, Honolulu, Hawaii

ALISON K. FLYNN

Veterinary Medical Center, University of Florida, Gainesville, Florida

6.0. CHAPTER SUMMARY AND OBJECTIVES

6.0.1. Summary

Along the information pipeline from DNA (a gene) to the production of a protein, there are many steps where gene expression can be controlled. In eukaryotes, such as plants, transcriptional control is considered the major form of gene regulation. Because of its importance, transcriptional regulation has been the best studied and probably the most manipulated. However, it is becoming clearer each day that posttranscriptional mechanisms of gene regulation are critical because levels of transcription are not always well correlated with functional protein levels. Additionally, as the area of proteomics advances, and as we move from genetically engineering plants to improve their performance or enhance their utility in a traditional agricultural setting, to using plants as biofactories to produce proteins, posttranslational regulation will gain in importance.

6.0.2. Discussion Questions

1. What are the differences between DNA and RNA?
2. Describe the main parts of a gene and their functions.
3. How important are *cis*-regulatory elements and *trans*-acting factors in gene regulation?
4. What are the control points that can regulate gene expression?

6.1. THE GENE

6.1.1. DNA Coding for a Protein via the Gene

From Chapter 2, we saw that there are several definitions of a gene. In this chapter, *gene* means a specific segment of DNA, including its regulatory regions, that code for a protein. In this chapter, we describe the *central dogma* of genetics, which involves information flow from DNA to RNA via

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

transcription in the nucleus, followed by RNA transport into the cytoplasm, where it is translated into protein. Let's first look at DNA. What exactly is DNA? DNA, or *deoxyribonucleic acid*, is simply a chemical, a double-stranded, helical polynucleotide, to be specific. However, in the proper biological context, this chemical determines such traits as the color of a petunia petal, the scent of a citrus blossom, the sweetness of a corn kernel, the strength of a cotton fiber, and the yield of a wheat head in the face of biotic and abiotic stress. The majority of a plant's DNA is found within the nucleus of each cell. Specific segments of the nuclear DNA, called *genes*, contain all the information required for the cell to make proteins (polypeptides) that are responsible for traits. Each protein-coding gene codes for a particular polypeptide, which is composed of a unique linear arrangement of amino acids as determined by the gene sequence.

6.1.2. DNA as a Polynucleotide

Before describing how the DNA of a gene can lead to the production of a protein (gene expression), the chemical structure of DNA must be understood. DNA is composed of two strands of *deoxyribonucleotides* (sugar (*deoxyribose*) + phosphate + a nitrogenous *base*—(either adenine (A), guanine (G) (both are *purines*), cytosine (C), or thymine (T) (both are *pyrimidines*)) (Fig. 6.1). The two strands have a right-handed (clockwise) helical shape, the so-called double helix (Watson and Crick's model), with the sugars and phosphates forming the backbone (or outside), and the bases located in the center of the molecule (Fig. 6.2). It is important to note here that the phosphates of the DNA backbone are negatively charged, and this will allow proteins that have positively charged domains to bind to the DNA. The importance of such DNA–protein binding will be discussed later in this chapter in terms of controlling gene expression. The deoxyribonucleotides of each strand are paired through specific hydrogen bonding of their respective bases: A on one strand always pairs with T on the other via two hydrogen bonds, and G on one strand always pairs with C on the other via three hydrogen bonds. This hydrogen bonding keeps the two strands together. Knowing the sequence of only one of the strands will provide all the information required to make the other strand through this specific or complementary base-pairing mechanism. It is also sufficient information for scientists to deduce the sequence of the second strand. It is important to note that the strands have directionality, each has a 5′ end and a 3′ end, and when the DNA strands pair, they are said to be antiparallel (Fig. 6.1). Since the bases are what distinguish the nucleotides from one another, a gene sequence conventionally is written by listing the linear sequence of the bases of one strand (the *coding* strand; see the text) starting from the 5′ end and proceeding to the 3′ end.

6.2. DNA PACKAGING INTO EUKARYOTIC CHROMOSOMES

In a cell, the DNA described earlier is not "naked," but in association with proteins that together are packaged as chromosomes that can fit within the nucleus. Specifically, eukaryotic chromosomes are composed of DNA (2 nm in diameter) in association with histone and nonhistone proteins to form a nucleoprotein structure called *chromatin* (200 nm in diameter). Initially, the histones produce a complex with the DNA to form the first structural unit of chromatin called the *nucleosome*. The nucleosome consists of DNA wrapped twice around a core of eight histones to form a 10-nm fiber (Fig. 6.3). This fiber is further folded to result in a chromatin fiber. Chromatin is dynamic in terms of its ability to coil and uncoil during the lifecycle of a cell. Chromatin is in its most condensed or coiled form during mitosis, when it forms a metaphase chromosome (700 nm in diameter). Regulation of gene expression, as detailed later in the text, involves nucleosome uncoiling and this change in DNA conformation is termed *chromatin remodeling*. So, chromatin not only is necessary for packaging DNA to conveniently fit within the nucleus of a cell but also plays an important role in gene expression.

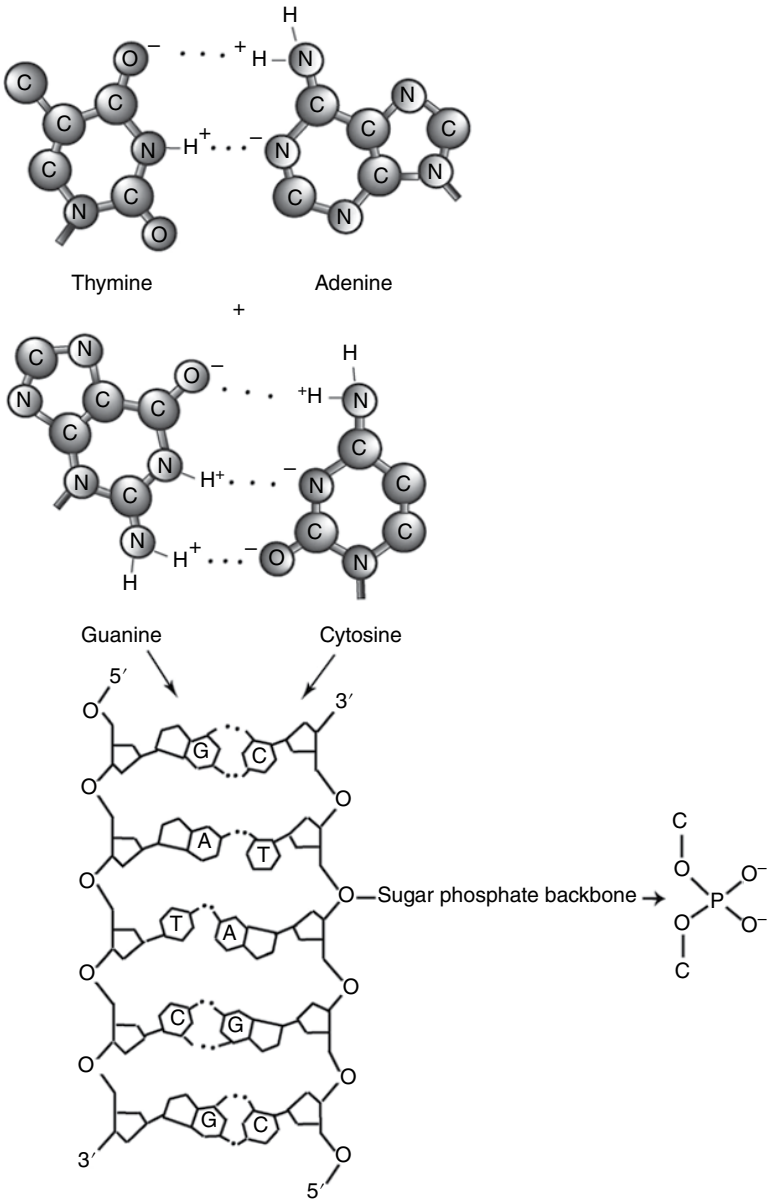


Figure 6.1. DNA structure. The purine nucleotides (adenine and guanine) form hydrogen bonds with the pyrimidine nucleotides (thymine and cytosine, respectively). Nucleotides are strung together by a sugar phosphate backbone that has an antiparallel orientation (5′–3′) to the complementary base pairs.

6.3. TRANSCRIPTION

6.3.1. Transcription of DNA to Produce Messenger Ribonucleic Acid

How does the information contained in a protein-coding gene on a chromosome within the nucleus lead to the formation of a polypeptide in the cytoplasm? The key is that the DNA of a gene does not directly participate in the synthesis of a polypeptide. The gene's information or message is faithfully carried by another molecule out of the nucleus and into the cytoplasm. The first step in this information flow from DNA to polypeptide is to synthesize this "messenger" from the gene in a process called

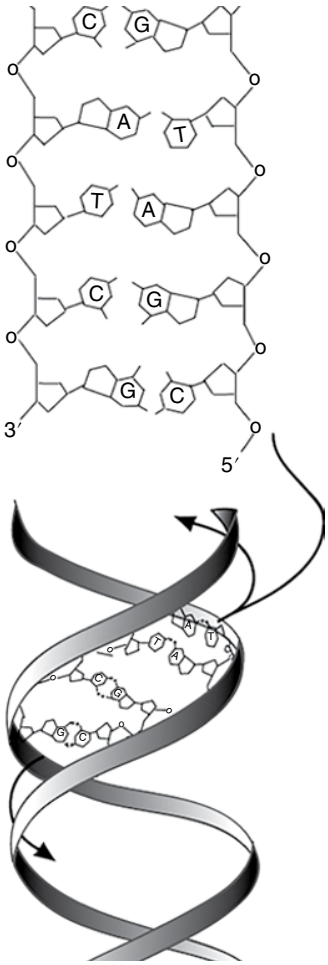


Figure 6.2. The two antiparallel chains of nucleotides strung together by a negatively charged sugar phosphate backbone form a right-handed double helix with the base pairs in the center of the helix.

transcription. The transcribed messenger molecule, also referred to as a *transcript*, is another polynucleotide aptly named *messenger ribonucleic acid* (mRNA). Like DNA, mRNA is composed of nucleotides that are assembled in a $5' \rightarrow 3'$ direction; however, mRNA is made up of *ribonucleotides*, because its sugar is a *ribose*. mRNA also differs from DNA in that it is a single-stranded molecule and, in place of T, it has another base, uracil (U), which can form a complementary base pair with A. Only one DNA strand of a gene is used as a "template" during transcription to create the mRNA. The order or linear sequence of bases in this DNA *template* strand ($3' \rightarrow 5'$) determines the sequence of the mRNA ($5' \rightarrow 3'$) because transcription works through complementary base pairing. Consequently, the mRNA made is a complement of the DNA template strand of the gene and an exact copy of the other DNA strand of the gene (the *coding* strand) except for having a U where a T would be located (Fig. 6.4).

Transcription is carried out by the enzyme *RNA polymerase II* (RNAP II) in eukaryotes such as plants. RNAP II does not act alone. Its binding and activity are controlled by both DNA sequences located within the gene (*cis-regulatory region*) and by proteins (*transacting factors*) called *transcription factors*, which can be general in helping transcribe many genes or specific to one or a few genes. Gaining a better understanding of the roles that *cis*-regulatory regions and transcription factors play in gene regulation is an active area of current research. The *general transcription factors* (GTFs) are necessary for RNAP II to transcribe DNA. The specific transcription factors affect the efficiency or the rate of RNAP II transcription for specific genes.

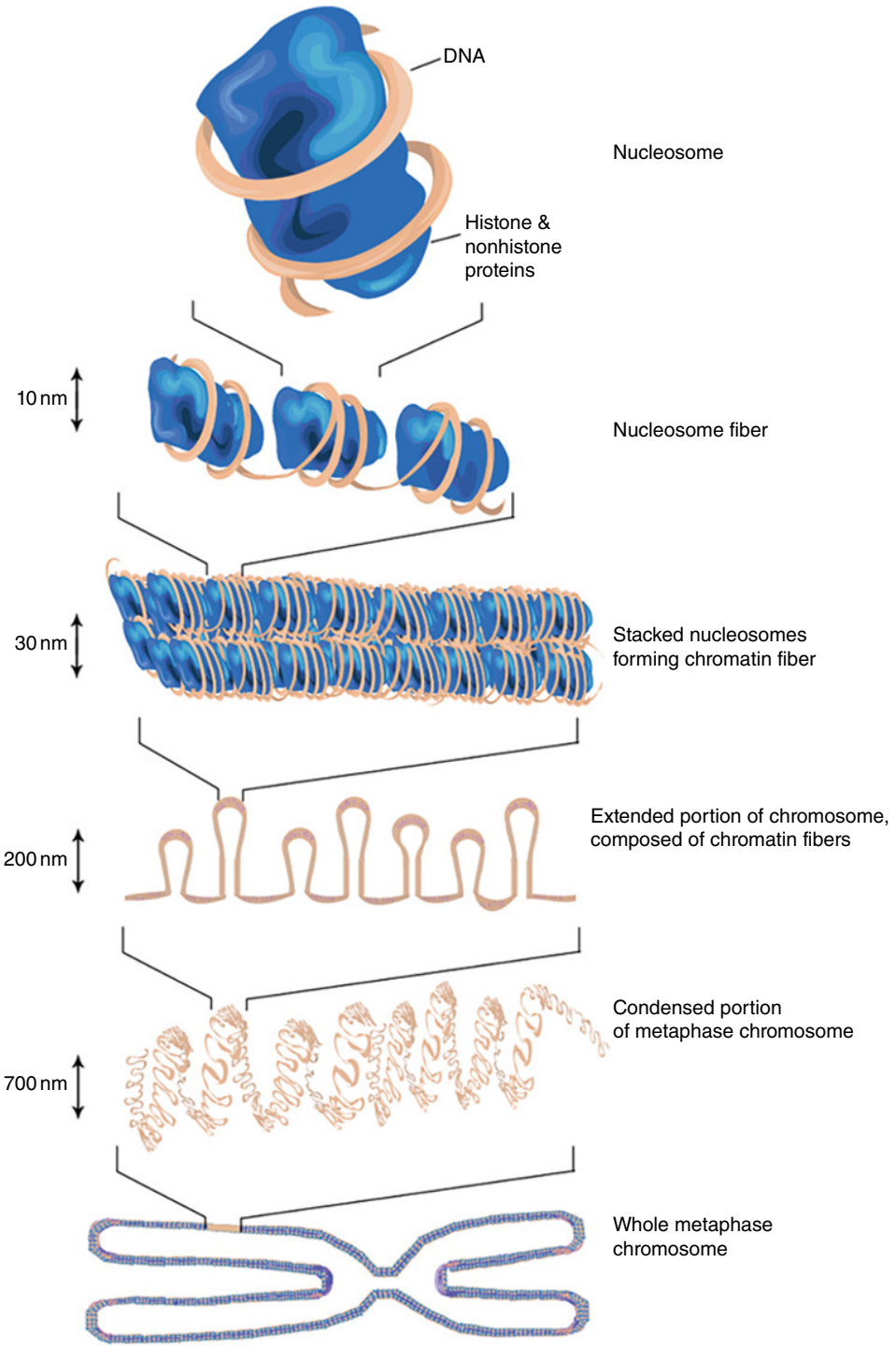


Figure 6.3. Chromatin structure. The different levels of chromatin structure are shown. The basic building block consists of the nucleosome–DNA wrapped around a core of histone and nonhistone proteins. Nucleosomes are strung together by strands of DNA, which are densely packed to create chromatin. Tight winding of chromatin creates the metaphase chromosome.

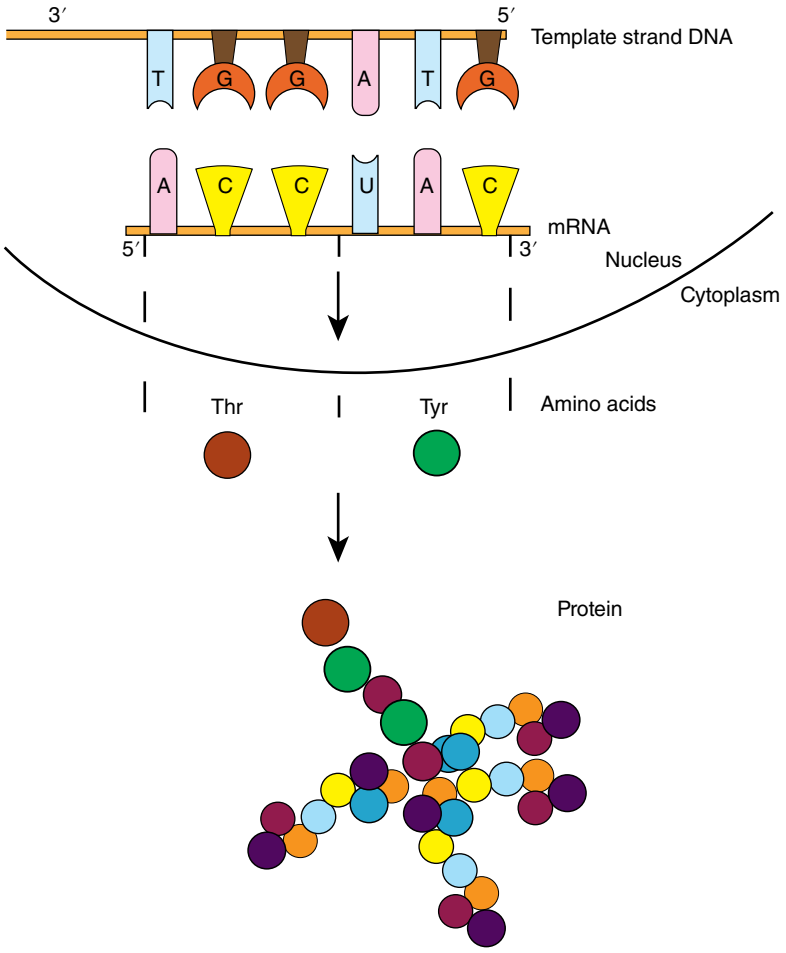


Figure 6.4. The central dogma: DNA is transcribed to RNA in the cell nucleus. RNA is translated to protein in the cell cytoplasm. (*See insert for color representation of the figure.*)

The *cis*-regulatory region controlling transcription by RNAP II, called the *promoter*, is located at a gene's 5′ end (using the coding strand as a reference). The promoter is composed of a core promoter plus other promoter elements that help define when and where a gene is transcribed. The core promoter element is where RNAP II and the GTFs bind to begin transcription. The *transcription start site* (the gene location where the first ribonucleotide of the RNA being synthesized will base pair) is designated as the +1 site (i.e., the first base in the transcript), so the gene promoter is therefore located upstream of (or before) the +1 site and its nucleotides are given negative numbers, whereas all nucleotides after the +1 site are positive sequential numbers (Fig. 6.5). As will be detailed later, the actual protein-coding portion of the gene will begin with an ATG sequence (AUG in the mRNA), but the +1 site is generally well upstream or in front of that sequence. Therefore, the portion of the gene from the +1 site up until the ATG sequence is termed the 5′ *untranslated region* (5′UTR: this sequence is located in the gene and in the transcribed mRNA, but it does not get read for translation). Similarly, at the end of a gene, there is also a portion that is transcribed into mRNA, but is not translated, and that is termed the *3′ untranslated region* (3′UTR).

A core promoter element found in most eukaryotic genes consists of a *consensus sequence* (the bases most often found at certain positions that have been conserved throughout evolution) located at approximately −25 to −30 called the *TATA box* or the *Goldberg–Hogness box* (Goldberg 1979). It is called TATA because the bases T and A are prominent. Initially, RNAP II and the GTFs are

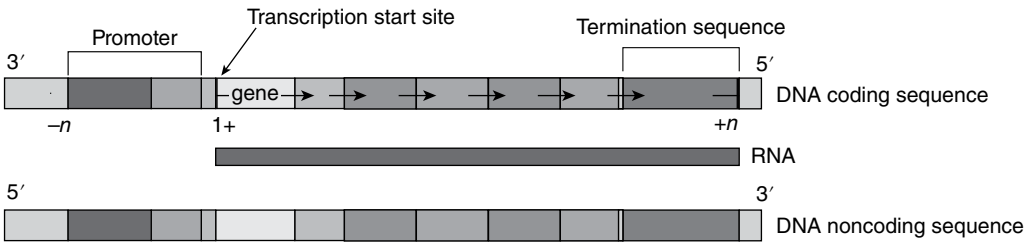


Figure 6.5. The structure of the promoter in relation to the gene and termination sequence. The transcription start site is designated by 1^+, and DNA sequences that are downstream or toward the 3′ end of the DNA strand are described by a negative number. In contrast, nucleotide sequences upstream (toward the 5′ end of the DNA strand) are described with a positive number.

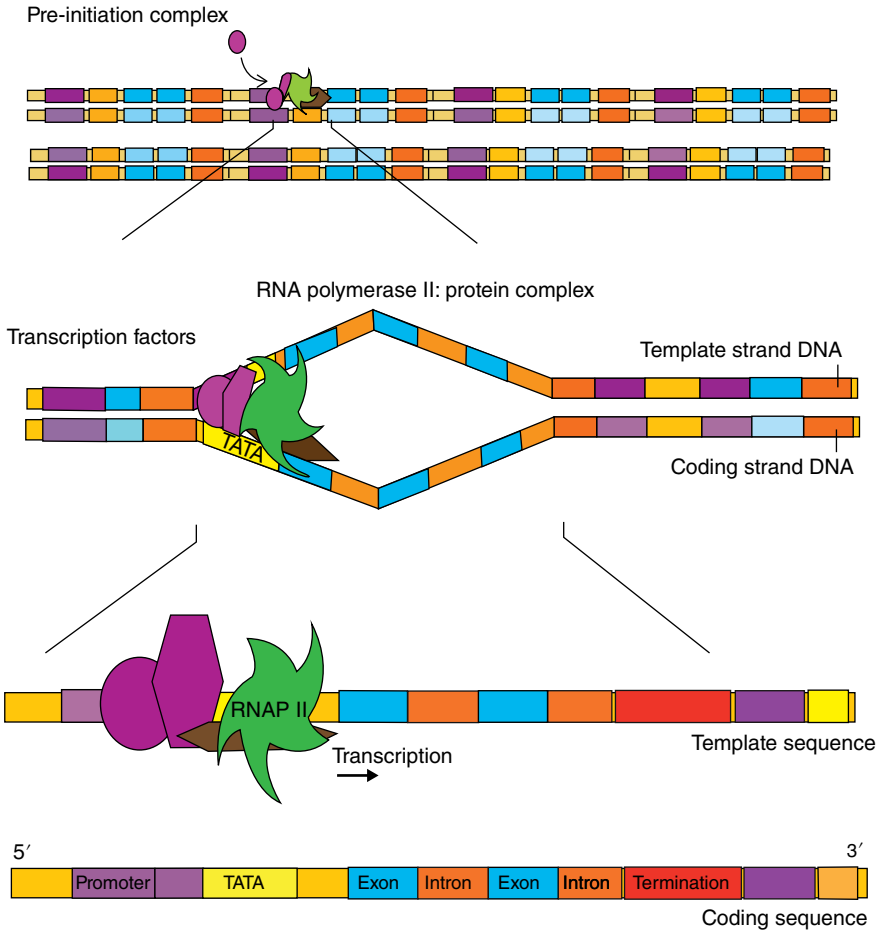


Figure 6.6. Overview of the early steps of transcription. A preinitiation complex is formed by a complex of transcription factors and RNA polymerase II (RNAP II). Association of the preinitiation complex with the start sequence (TATA) of the coding strand of DNA causes a conformation change and hydrogen bond breakage. This causes the DNA strands to separate so that transcription can proceed. (*See insert for color representation of the figure.*)

bound to the core promoter element in an inactive state called the *preinitiation complex* (PIC). Then 11–15 base pairs of the gene around the transcription start site break their bonds, thereby changing the DNA conformation into an open complex, and the template strand of the promoter becomes located in the active site of RNAP II to initiate transcription at a basal level (Fig. 6.6).

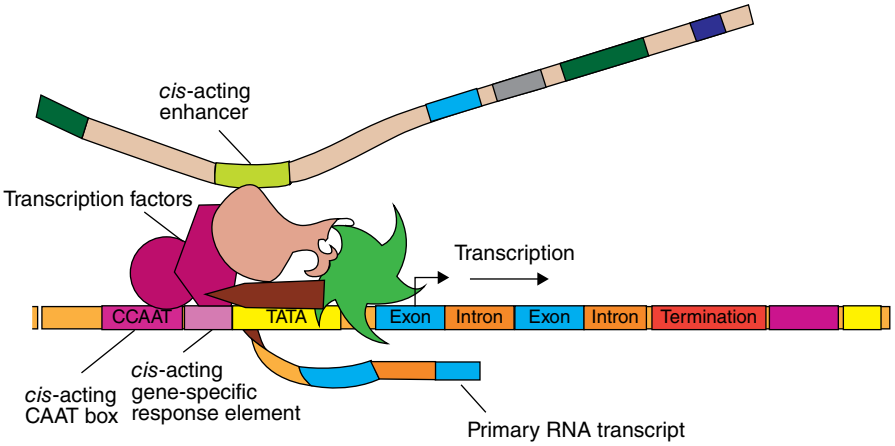


Figure 6.7. Regulation of transcription. The *cis*-acting elements are segments of DNA that regulate transcription; these segments may be adjacent to the gene such as the promoter (CAAT box) and the *cis*-acting gene-specific response elements, or they may be distant to the gene such as enhancers. The *trans*-acting elements are transcription factors and other regulatory proteins that may associate with the promoter, other proteins, or both. (*See insert for color representation of the figure.*)

Promoter elements that are not required for transcription initiation, but influence the level, rate, timing, or tissue specificity of transcription are the *CAAT box* (CCAAT), and gene-specific response elements. The CAAT box is generally located at sites −70 to −80. The gene-specific response elements vary in their sequence and location within the promoter. A third type of *cis*-regulatory element is an *enhancer*, the location of which varies from gene to gene. Unlike a promoter element, an enhancer can function even at long distances (>1 kb) upstream or downstream of the transcriptional start site (Khoury and Gruss 1983), and its orientation can be inverted without losing its function. The CAAT box, gene-specific response elements, and enhancers carry out their functions by binding specific transcription factors (Fig. 6.7).

6.3.2. Transcription Factors

Transcription factors are regulatory proteins that bind to DNA and to other regulatory proteins to affect gene expression, as described earlier. Thus, there are transcription factor genes whose expression affects the regulation of other genes. To carry out their functions, they generally have specific portions, or domains. There are two main domains in transcription factors, a *DNA-binding domain* and a *trans-acting domain*. The *DNA-binding domain* does just that; it allows the transcription factor to bind directly to a DNA *cis*-regulatory element. DNA-binding domains are characterized by specific structures or motifs. For example, some DNA-binding domains have a helix–turn–helix motif, a zinc finger motif, or a leucine zipper motif. The *trans-acting domain* of a transcription factor allows it to bind to RNAP II or to other transcription factors, thus allowing protein–protein interactions. So, with two such domains, a transcription factor can simultaneously bind DNA and other transcription factors or RNAP II to regulate gene transcription (Fig. 6.8).

6.3.3. Coordinated Regulation of Gene Expression

Eukaryotes can coordinately express subsets of many different genes in response to particular biotic and abiotic signals because those genes will contain common *cis*-regulatory or response elements in their promoters or enhancers that allow them to recognize the same signals. These elements have a consensus sequence that can bind specific transcription factors allowing for transcription of those

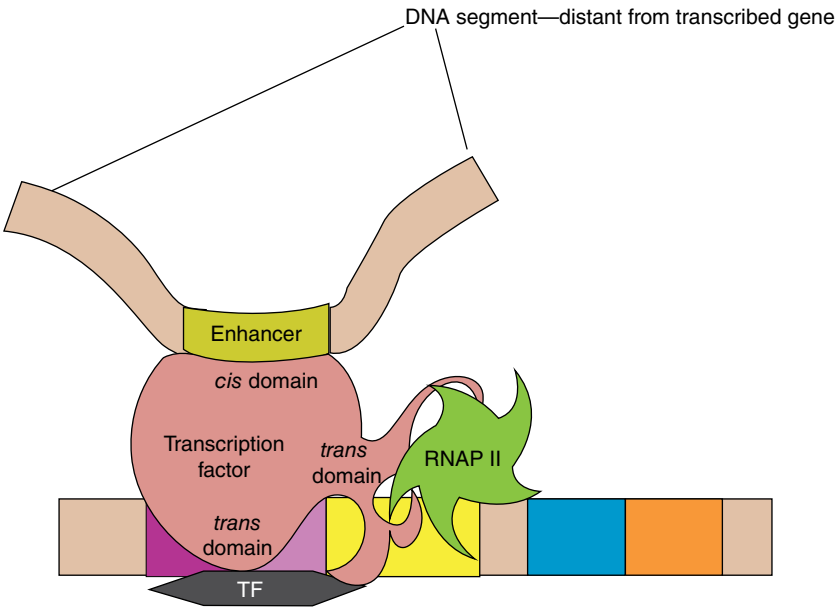


Figure 6.8. Transcription factors structure and function. Transcription factors may have domains that bind *cis*-acting elements such as enhancers, and domains that also bind *trans*-acting elements such as RNA polymerase (RNAP II) and other transcription factors. (*See insert for color representation of the figure.*)

genes. A gene may also contain several different response elements allowing it to be expressed following a number of stimuli. For example, the CBF transcription factors of *Arabidopsis* (Gilmour et al. 1998) can bind to the cold- and dehydration-responsive *cis*-regulatory element called C-repeat/dehydration-responsive element (CRT/DRE) (Baker et al. 1994; Yamaguchi-Shinozaki and Shinozaki 1994) that is found in the promoters of many cold- and dehydration-responsive genes of *Arabidopsis*. So, following cold or water-stress stimuli, those genes containing the CRT/DRE-responsive element will be transcribed and provide *Arabidopsis* with increased tolerance to freezing, as well as drought.

6.3.4. Chromatin as an Important Regulator of Transcription

DNA wrapped around histones and coiled to produce chromatin is not accessible for transcription. It is not physically possible for RNAP II to make contact with the DNA for transcription. Chromatin remodeling, as mentioned earlier, is required to allow the appropriate regions of a gene to bind transcription factors and RNAP II for transcription. This remodeling "opens up" the DNA to make it accessible to RNAP II and transcription factors. Following remodeling, the promoter region no longer contains histones, thereby making the *cis*-regulatory elements free to bind to the necessary transcription factors and RNAP II to begin transcription. Chromatin remodeling is done by various multiprotein complexes that have ATP-ase activity (use energy) to bind directly to particular regions of the DNA to move the nucleosomes to a new position to expose the DNA for transcription (Vignali et al. 2000).

Chromatin structure also can be changed through the covalent addition of acetyl groups (CH_3CO) to the histones of the nucleosome. When the acetyl groups are added to the histone tails, they are no longer positively charged, and consequently the negatively charged DNA can disengage from them. The acetyl groups are added by enzymes called *histone acetyltransferases* (HATs). It is known that certain transcription factors have acetyltransferase activity or can recruit these enzymes to the DNA, thereby altering chromatin structure and allowing for transcription. Chromatin structure can be

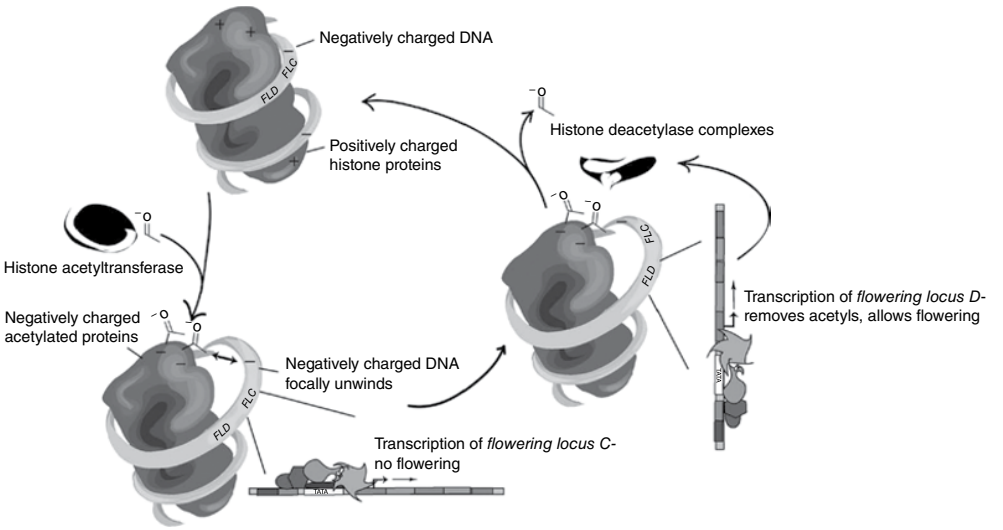


Figure 6.9. Control of transcription by chromatin remodeling. Genetic regulation of flowering in *Arabidopsis*. If the gene *flowering locus C* (*FLC*) is expressed, flowering does not occur. However, if the *flowering locus D* (*FLD*) gene is active, it produces a deacetylase that removes acetyls from histones around *FLC*. Consequently, transcription of *FLC* cannot occur because of the restoration of chromatin structure, and silencing of *FLC* allows flowering.

restored by histone deacetylase complexes (HDACs) that remove the acetyl groups from these histones. A good example of such regulation of gene expression can be found in the control of flowering in *Arabidopsis*. If the gene *flowering locus C* (*FLC*) is expressed, flowering does not occur. However, if the *flowering locus D* (*FLD*) gene is active, it produces a deacetylase that removes acetyls from histones around *FLC*. Consequently, transcription of *FLC* cannot occur owing to the restoration of chromatin structure, and silencing of *FLC* allows flowering (Fig. 6.9) (He et al. 2003).

6.3.5. Regulation of Gene Expression by DNA Methylation

DNA methylation (–CH_3 groups attached to DNA of the promoter or coding region) is a major factor in regulating gene expression. There appears to be an inverse relationship between percent methylation and the degree of expression. *Hypo*methylation is associated with higher gene expression, whereas *hyper*methylation is associated with greater gene silencing. The most common methylated base in eukaryotic genomic DNA is 5-methyl-cytosine (m^5C).

Plants generally have higher levels of DNA methylation than do mammals. Also, in plants methylation occurs mainly in transposable elements and other repeat sequences. If a transposon is methylated, it will be inactive and not hop around the genome, but it can be activated if the methylation is removed. However, as in mammals, methylation of the cytosine on both strands of the CpG dinucleotide (linear sequence of cytosine followed by a guanine separated by a phosphate, to be distinguished from a cytosine base-paired to a guanine) is common in plants and is carried out by DNA methyltransferases such as MET1 in *Arabidopsis*. This enzyme is responsible for the maintenance of global genomic methylation. Plants mutant for MET1 have significantly lower levels of methylation and show late flowering phenotypes (Kankel et al. 2003). Also, transgenes that are genetically engineered into plants and become highly methylated are not expressed. However, if these plants have a defective MET1, these transgenes will no longer be silenced. Plants also have methylation of CpNpG trinucleotides ("N" can be any of the four DNA bases) and asymmetric CpNpN trinucleotide sites that are performed by specific enzymes that are unique to plants such as *chromomethylases* (CMTs) and *domain-rearranged methylases* (DRMs). The CMTs appear to be

involved in maintaining methylation of sites that are heavily methylated to keep them silenced. The DRMs function in RNA-directed DNA methylation by somehow recognizing *small interference RNA* (siRNA—these RNAs are usually 20–25 nucleotides long and inhibit expression of specific genes) and then methylating the appropriate DNA sequences. Additionally, it has been shown that chromatin-remodeling factors, as described earlier, can be necessary for maintaining methylation.

6.3.6. RNA-Directed Gene Silencing by Small RNAs

Small (generally 21–25 nucleotides in length), non-coding RNAs such as siRNAs mentioned earlier, as well as microRNAs (miRNAs), are important as negative regulators of gene expression. The siRNAs come from long, double-stranded RNAs or hairpin RNAs that are originally from viruses, repetitive, or transposon-rich sequences. The miRNAs are created from single-stranded transcripts called pri-miRNAs produced from endogenous MIR genes in the cell. Both types of small RNAs are critical components of what is called a RNA-induced silencing complex (RISC). Each RISC targets mRNAs in the cell that have complementary sequences to the small RNA it contains and that results in the targeted mRNA being cleaved, degraded, or not translated. This type of posttranscriptional regulation has been termed RNA interference (RNAi; Fire et al. 1998).

RNAi has been widely used for plant improvement research to develop crops that are resistant to pathogens and insect pests, have improved nutrition, and are more tolerant to abiotic stress. RNAi is also a valuable tool for basic functional genomic studies utilizing the power of reverse genetics.

In addition to posttranscriptional regulation, these small RNAs also can control chromatin structure by modifying histone deacetylases and influencing DNA methylation by modulating DNA methyltransferases. Consequently, small RNAs are essential players in the mechanisms of epigenetics, heritable changes in gene expression that are not the result of mutations in the DNA coding sequence.

6.3.7. Processing to Produce Mature mRNA

Controlling transcription is one of the most important ways to alter gene expression for biotechnology applications. Many of the mechanisms that plants possess to regulate the transcription of DNA to mRNA have been introduced earlier. Promoters, transcription factors, chromatin remodeling, and DNA methylation are all crucial for transcriptional control. However, transcription is only the first step in gene regulation. The mRNA that is made through transcription is not mature and is termed a *pre-mRNA* or a *heterogeneous nuclear RNA* (hnRNA). Before a gene transcript is transported out of the nucleus and into the cytoplasm where it will ultimately be translated into protein, it must be processed in several ways: *5′ capping*, *3′ polyadenylation* (*polyA tail*), and *splicing* out of *introns* and putting together of *exons* (Fig. 6.10). The first processing step, occurring when approximately 20–30 ribonucleotides of the hnRNA have been made, is the addition of a 7-methylguanosine to the 5′ end of the transcript. This cap structure may play a role in mRNA stability by physically protecting the mRNA from $5' \rightarrow 3'$ *exonucleases*, types of *RNAses*, once it is in the cytoplasm. Most eukaryotic gene protein-coding regions are interrupted by non-protein-coding sequences (introns) that are removed from the hnRNA, so they are not found in mature mRNA. The hnRNA has consensus sequences at the exon–intron junctions (marked by "/") that are required for proper splicing out of the introns. The 5′ exon–intron junction consensus sequence is AG/GURAGU, and the 3′ exon–intron junction consensus sequence is YAG/RNNN (Y, pyrimidine; R, purine; and N, either purine or pyrimidine). Also, about 100 nucleotides upstream of the 3′ exon–intron junction, there is a branchpoint conserved sequence. A *splicesome* composed of small nuclear RNAs (snRNAs—these RNAs are 100–300 bases in length) and various proteins forms over the intron and helps in the splicing process. Most mRNAs contain a polyadenylated 3′ end consisting of 200 A residues. This polyA tail acts as a protective buffer against RNAses that could digest the mRNA from the 3′ end, and thus stabilizes the molecule. Approximately 10–30 bp upstream of this polyA tail is the invariant hexamer sequence AAUAAA (Fig. 6.10).

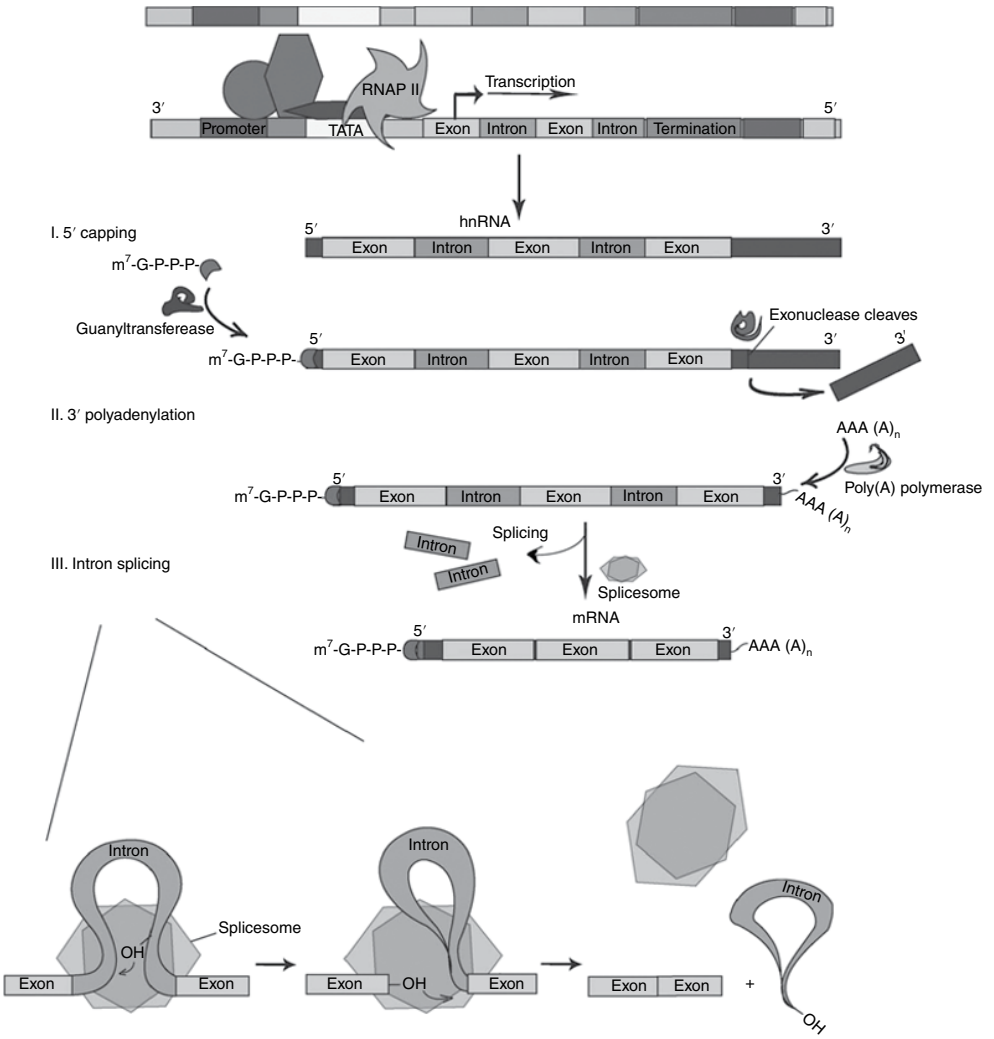


Figure 6.10. Overview of mRNA processing. Three steps of processing must occur prior to export of the mature mRNA out of the nucleus: 5′ capping, 3′ polyadenylation, and intron splicing.

6.4. TRANSLATION

How does the information in the mRNA result in the synthesis of a polypeptide? Multiple cellular players are involved in the synthesis of a polypeptide. First, the structure of a polypeptide needs to be understood. Polypeptides are made up of a linear sequence of amino acids. There are 20 common types of amino acids (Table 6.1); and to form a polypeptide, amino acids are joined together in a chain by peptide bonds (Fig. 6.11). Proteins can be composed of either a single polypeptide or multiple polypeptide chains that are the same or different in amino acid sequence (Fig. 6.11).

Once the mRNA is transported out of the nucleus, it must be properly "read" or *translated* by *ribosomes* in order to produce a polypeptide. But how many nucleotides of the mRNA are needed to code for one amino acid? Three consecutive nucleotides, called a *codon*, are required to be read to specify one amino acid. This code is *nonoverlapping*, meaning that once a triplet is read, the cellular machinery reads the next three nucleotides and so on in a linear fashion. Therefore, within a given *reading frame* (there are three possible reading frames; see Fig. 6.12), a nucleotide cannot be present in more than one codon. Since there are four nucleotide possibilities (A, G, C, or U) at each of the

TABLE 6.1. The 20 Amino Acids Commonly Found in Proteins

Amino acid	Three-letter abbreviation	One-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamate	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Try	W
Tyrosine	Tyr	Y
Valine	Val	V

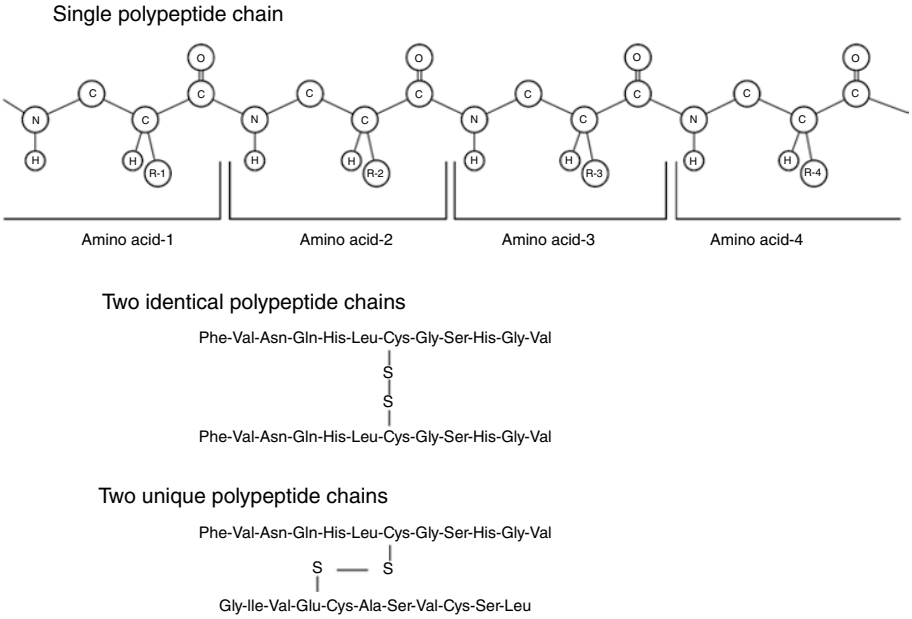


Figure 6.11. Polypeptide structure. The building block of a polypeptide is the peptide bond formed between amino acids. Peptide bonds connect amino acids to create a polypeptide chain. Proteins are formed through the association of individual polypeptide chains that may be identical to each other or unique in sequence.

three codon positions, there are $4 \times 4 \times 4 = 64$ different combinations or codons (Table 6.2). A codon is written in the $5' \rightarrow 3'$ direction as it would be read on the mRNA molecule. Since there are more codons than amino acids, some codons actually specify the same amino acid, and so the code is considered to be degenerate in that regard. Three codons (UAA, UAG, and UGA) do not code for any

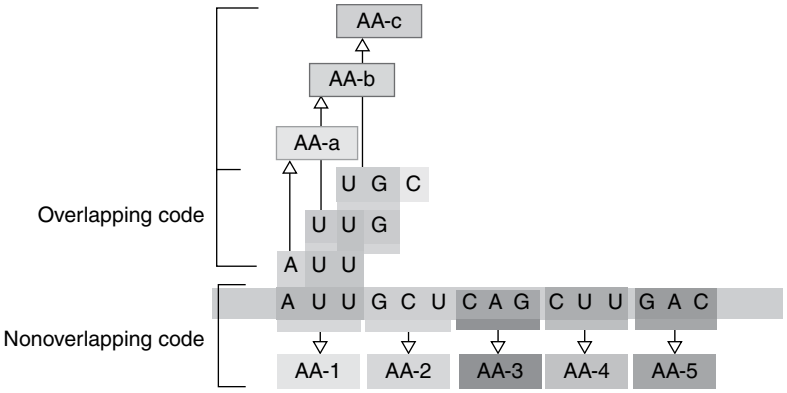


Figure 6.12. The genetic code gives rise to either overlapping or non-overlapping reading sequences. A codon consists of three consecutive nucleotides that code for an amino acid. The nucleotides in a codon may give rise to multiple amino acids depending on the reading frame.

TABLE 6.2. The Genetic Code—mRNA Codons and Amino Acids Encoded

First base	Second base								Second base
	U		C		A		G		
	UUU	Phe	UCU		UAU	Tyr	UGU	Cys	U
U	*UUC*		UCC	Ser	*UAC*		*UGC*		C
	UUA	Leu	UCA		UAA	Stop	UGA	Stop	A
	UUG		UCG		UAG		UGG	Trp	G
	CUU		CCU		CAU	His	CGU		U
C	CUC	Leu	CCC	Pro	*CAC*	Gln	CGC	Arg	C
	CUA		CCA		CAA		CGA		A
	CUG		CCG		CAG		CGG		G
	AUU		ACU		AAU	Asn	AGU	Ser	U
A	*AUC*	Ile	ACC	Thr	*AAC*		*AGC*		C
	AUA		ACA		AAA	Lys	AGA	Arg	A
	AUG	Met	ACG		AAG		AGG		G
	GUU		GCU		GAU	Asp	GGU		U
G	GUC	Val	GCC	Ala	*GAC*		GGC	Gly	C
	GUA		GCA		GAA	Glu	GGA		A
	GUG		GCG		GAG		GGG		G

The codons are written in the $5' \rightarrow 3'$ direction.

amino acid. These are *stop codons*, and when any one is read, it signals the cellular machinery to stop translation.

Since an mRNA is a long molecule containing many nucleotides, where does the translational machinery start looking to begin reading each codon? The first codon read is called the *initiation* (or *start*) *codon*, and it is usually AUG that codes for methionine (earlier we mentioned that the protein-coding portion of the gene (DNA) began with ATG). In eukaryotes, the initiation codon is surrounded by a consensus sequence termed the *Kozak sequence* (ACC*AUG*G) (Kozak 1986, 1987), which indicates to the translational machinery to begin translation with this codon. If this sequence is not present, this codon will be missed and the cellular machinery will continue to scan down the mRNA until it finds a suitable initiation codon, if present. As mentioned earlier, three different reading

frames are possible. The start codon defines what the correct reading frame will be for any particular gene. As you will see later, this is an important consideration for biotechnology.

6.4.1. Initiation of Translation

Translation of the mRNA is done in connection with organelles called *ribosomes* and another type of RNA termed *transfer RNA* (tRNA). In eukaryotes, ribosomes are complex and composed of two subunits: large and small. The large subunit contains three types of ribosomal RNAs (rRNAs) (28S rRNA, 5S rRNA, and 5.8S rRNA), along with 49 proteins. The small subunit contains the 18S rRNA and 33 proteins. A ribosome will bind to the 5′ end of the mRNA and move down toward the 3′ end as translation proceeds. Specifically, starting at the 5′ cap of the mRNA, the small subunit of the ribosome along with initiation factors will bind and move down the mRNA until it encounters the proper initiation codon. Then the correct amino acid (the initiation codon codes for methionine; therefore, methionine is always the first amino acid in the initial polypeptide) is brought to it via a tRNA molecule and combines with additional factors to form an initiation complex. The tRNA molecule is said to be "charged" when it carries an amino acid. The charged tRNA molecule recognizes the codon through complementary base pairing with a region of it called an *anticodon* (Fig. 6.13).

6.4.2. Elongation Phase of Translation

Now polypeptide synthesis takes place with amino acids joining together as successive codons are read in the elongation phase of translation. Before elongation can occur, the large ribosomal subunit joins to create a complete ribosome. The ribosome now has three sites that can accommodate a tRNA molecule: a peptidyl (P), an aminoacyl (A), and an exit (E) site. The initiator tRNA occupies the P site of the ribosome, which is positioned over the initiator AUG codon and is adjacent to the A site, which at this stage is available and is over the next codon to be read. Then the appropriately charged tRNA for this next codon in the A site enters it, and its anticodon pairs with the codon. A peptide bond then forms between the amino acids that are attached to the tRNAs in the P and A sites. Now the initiator amino acid is released from its tRNA and the ribosome moves down the mRNA or translocates to position the growing polypeptide in the P site and free the A site, which once again positions over the next codon to be translated. The initiator tRNA that no longer is charged is in the E site, and it is then free to leave the ribosome and become charged again. This elongation cycle is repeated until the entire polypeptide chain is made.

6.4.3. Translation Termination

Polypeptide synthesis is over when the ribosome encounters a stop codon in its A site. Since no tRNAs can base pair with these stop codons, proteins called "release factors" bind to the ribosome instead. These release factors allow the polypetide chain to be released from the P site as well as the mRNA to no longer bind to the ribosome. The ribosome also splits into its two subunits.

6.5. PROTEIN POSTRANSLATIONAL MODIFICATION

Following translation, polypeptides can be modified in a number of ways before they are fully functional, and in fact, different organisms modify proteins in different ways that can have biological significance. The initiator amino acid, methionine, can be changed or removed. More amino acids can be added, or the polypeptide can be "trimmed" by removing amino acids. Also, amino acids can be modified by the addition of carbohydrate sidechains, phosphates, methyl groups, or conjugated with metals. These modifications can significantly alter the function of proteins, and subsequently

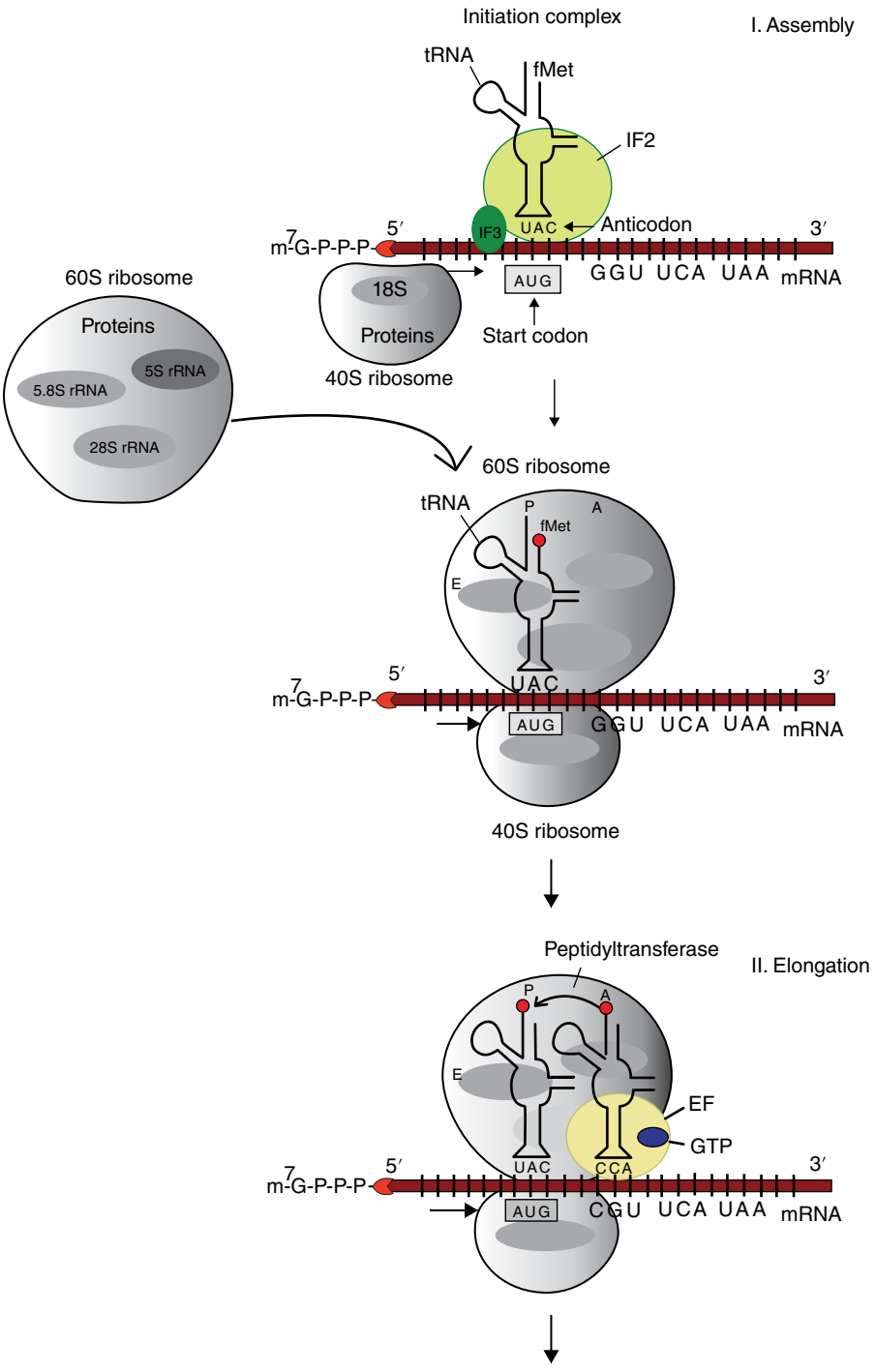


Figure 6.13. Overview of translation showing the structure of tRNA, 60S and 40S ribosomal subunits. The three steps of translation are shown: ribosome assembly, elongation of the polypeptide chain, and termination.

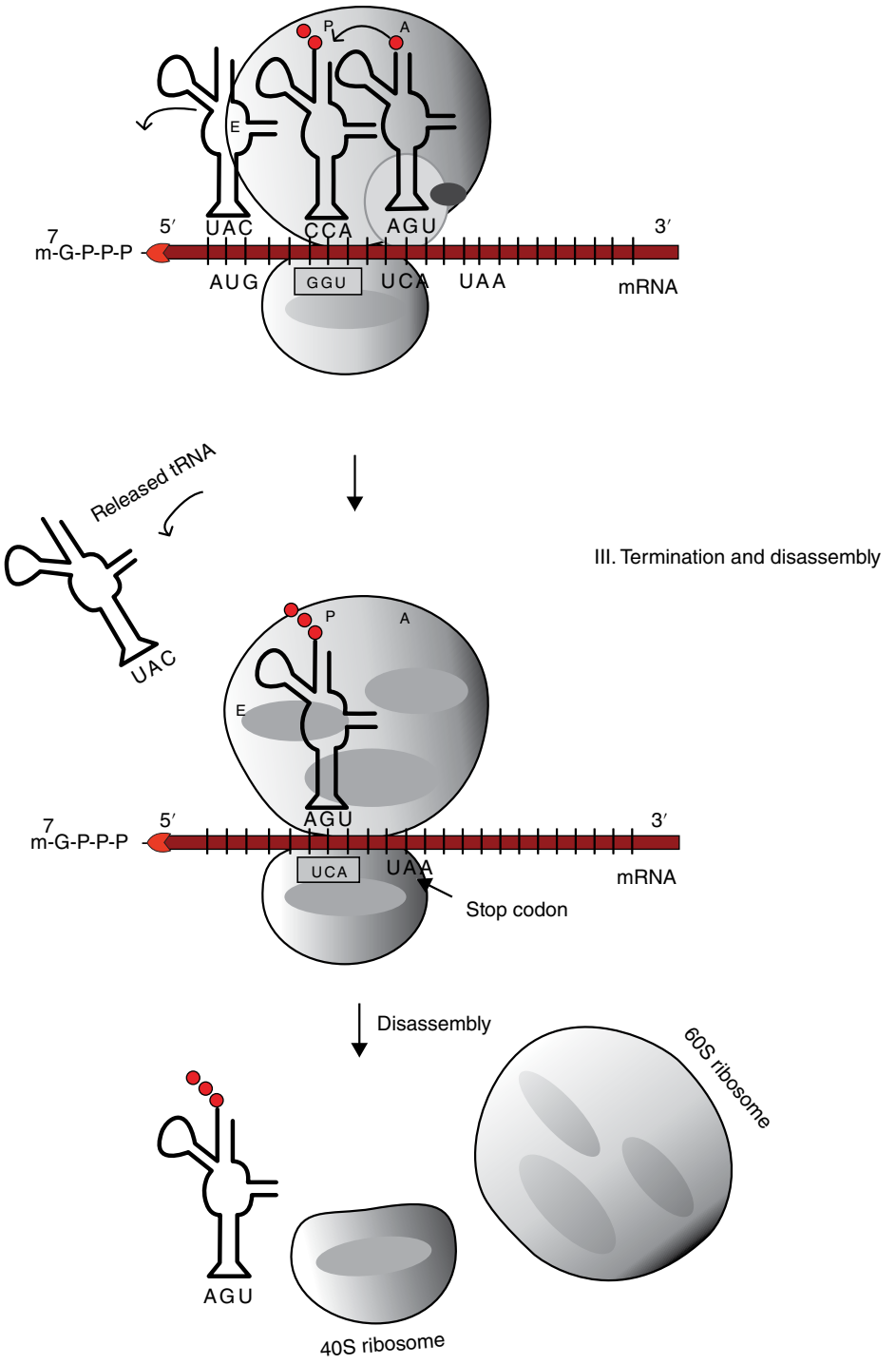


Figure 6.13. (Continued)

control cellular function. For example, phosphorylation is an important mechanism for controlling intracellular signaling. In order to be a functional protein, polypeptides also must be appropriately folded into a three-dimensional conformation, which can occur either spontaneously or under the direction of molecular "chaperones." As mentioned earlier, some proteins are composed of a single polypeptide, whereas others are multimeric, composed of one or more additional polypeptides that form the complete protein. Posttranslational modifications can fundamentally alter gene expression by changing protein function, allowing the cell to rapidly respond to variable internal and external stimuli. Understanding how to control posttranslational regulation is becoming increasingly valuable as we engineer some plants to be protein production factories, accumulating high levels of desirable, functional proteins for numerous applications ranging from industrial to medical.

LIFE BOX 6.1. MAARTEN CHRISPEELS

Maarten Chrispeels, Emeritus Distinguished Professor of Biology; University of California, San Diego, Member of the National Academy of Sciences.



Maarten Chrispeels on a bicycle with one of his students in Wuhan, China (October 2014). Courtesy of Maarten Chrispeels.

I was born in 1938 in a small Flemish town not far from Brussels, Belgium, and after an uneventful youth and a solid classical education with three foreign languages and 6 years of Latin, I enrolled in the College of Agriculture in Ghent. I wanted to become a biochemist. In the fall of 1960, after graduation, I found myself on the Mauretania, sailing for America with a Fulbright travel fellowship and a fellowship from the University of Illinois to start graduate studies in the Department of Agronomy. My Ph.D. research (with John B. Hanson) and postdoctoral work (with Joe Varner) was in plant cell biology. A couple of papers in plant physiology landed me a job as an assistant professor of biology at the then newly founded University of California San Diego (UCSD). Upon arriving there, I switched from studying α-amylase secretion by barley aleurone cells to the biosynthesis and secretion of hydroxyproline-rich cell wall glycoproteins. We found that these proteins move from the endoplasmic reticulum (ER) to the Golgi apparatus where glycosylation of hydroxyproline residues takes place. After a sabbatical leave in England, I switched to study the synthesis and intracellular transport of proteases to the protein storage vacuoles (PSVs) in seedling cotyledons. We made use of antibodies—quite a novelty at the time—to demonstrate by immuno-electron microscopy that the protease that digests storage proteins is in the ER before it arrives in the PSVs. It then occurred to me that if I wanted to study protein transport to PSVs I should be looking at developing seeds and not at germinating seeds, because seed development is characterized by massive protein synthesis and transport to vacuoles. So, I switched again and started working on the synthesis, posttranslational modification, and transport of storage glycoproteins and lectins in developing bean seeds. About that time, others invented gene cloning and plant transformation and soon we had bean storage protein and lectin genes and were expressing them in

tobacco seeds to identify vacuolar targeting domains. Protein targeting became the major focus of the lab.

Wishing to expand my horizons a bit more, we purified bean PSV membranes (tonoplasts really) and cloned the gene for the most abundant tonoplast protein. This turned out to be a protein with six membrane-spanning domains. What could its function be? Not until we obtained a homolog from *Arabidopsis* was Christophe Maurel in the lab able to show that this new family of proteins constituted the plant aquaporins. Aquaporins had been described the year before by Peter Agre who later received the Nobel Prize for this work. From then on, we dropped the targeting work and studied aquaporins, as they were capturing the imagination of many plant physiologists.

Somewhere along the line, I had time to do other things. My former mentor Joe Varner was always heavily involved in "service to the profession" and I also accepted to become first Associate Editor and later Editor-in-Chief of *Plant Physiology*, an excellent journal, then in need of a physical and intellectual facelift. During my 8 years as editor-in-chief, I believe that I contributed to this facelift, and my successor Natasha Raikhel took the journal to new heights.

When, in 1978, the USDA created its first competitive grants program, I called my friend Joe Key who had just been named the program director and volunteered to come to Washington DC on short notice to put together a panel to evaluate grants in the area of "Genetic Mechanisms for Crop Improvement." He took me up on my offer, and a few weeks later I was working in DC having received a leave of absence from UCSD. While on sabbatical leave in Canberra, Australia, in 1990 I became involved in a biotech project. We had isolated the cDNA for α-amylase inhibitor from common bean, and with my friend T. J. Higgins we expressed this gene in developing pea seeds. Larry Murdock from Purdue University showed that the larvae of the pea bruchid, which normally burrow into dry pea seeds, starved to death on these transgenic pea seeds, presumably because the bruchid digestive amylases are inhibited by the bean inhibitor. In the field, the transgenic peas were completely resistant to the bruchids. The technology has not yet been commercialized in part because the inhibitor also inhibits mammalian and avian amylases.

At that time, there was no good textbook to help university teachers who wished to teach courses in plant biology with an applied or biotechnology focus, and I started work on the first edition of a text that in its second edition was called *Plants, Genes and Crop Biotechnology*. David Sadava and I put together a completely integrated textbook that had elements of plant physiology and biochemistry, human nutrition, plant breeding, human population changes and world food production, soils and plant nutrition, and biotechnology applications.

By the year 1997, 30 years after my arrival in San Diego, plant biology had grown from just three faculty members to about 15 in three different institutions—UCSD, The Salk Institute, and The Scripps Research Institute—and we founded the San Diego Center for Molecular Agriculture—a virtual center with a grandiose name—whose purpose it would be to enrich our own intellectual lives. Creating "a community of scholars," as we fondly call academia, is actually quite difficult and requires effort and commitment from all parties. For 10 years, I organized yearly symposia with speakers from abroad that brought together the entire San Diego plant biology community.

I closed my lab and officially retired in 2007 but kept on working. From 2008 until 2013, I was director of ScienceBridge at UCSD and organized outreach to high schools, helping teachers bring cutting-edge scientific experiments to their classrooms. Interested in agricultural biotechnology, I taught short courses in the subject in Belgium, Italy, and Mexico, and most recently in Wuhan, China. At the close of 2014, I decided to start working on the third edition of our crop biotechnology college textbook.

LIFE BOX 6.2. DAVID W. OW

David W. Ow, Principal Investigator and Director, Plant Gene Engineering Center, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China.



David Ow while visiting the Cotton Research Institute, Anyang, China (June 2007). Courtesy of David Ow.

Chance events shape a career. In spring 1977, while an undergraduate in Rich Calendar's lab, I heard good reviews about Hatch Echols' graduate seminar course on "Genetics and Society." An extensive list of topics for presentations was available; but as a lowly undergraduate, I got left with the topic of least interest to others: agriculture and society. I was so worried that I might not measure up to the graduate students that I spent countless hours at the library reading up the Green Revolution and the promising technologies in plant tissue culture, nitrogen fixation, and photosynthesis. For a 1-unit pass/fail course, it turned out to be more work than any of my other classes. Unexpectedly, after boning up on this topic, I actually got excited about genetics for agriculture.[1] After my presentation, Hatch Echols in his usual tie-dye tee shirt had a chat with me about plants, agriculture, and the third world. He saw that I might want some practical experience and suggested I see Renee Sung. So I ended up moonlighting in Renee's lab to learn plant tissue culture. When it came time for graduate work, however, I decided on bacterial genetics; plant cell culture work was too slow for my liking. Fred Ausubel's lab was cloning bacterial nitrogen fixation genes, and also doing petunia cell culture, so I ended up at Harvard.

The summer before grad school, I took a month off to visit the Orient. In 1978, China was off limits except for a 3-day tour to Guangzhou. On the train from Hong Kong, I sat next to a Mrs. Bogorad.[2] Apparently, Lawrence Bogorad just left for an official delegation to Beijing, and put her on the Guangzhou tour until they could join up again. I don't know whether meeting her had any relevance; but in Fall 1978, during my rotation in Fred's lab, Dr. Bogorad called me to a reception for a Chinese delegation. Being a first-year student, I was a bit nervous but managed a good exchange with the Chinese visitors, who were all quite aged as they had received their PhDs from the West before communism. When we parted, some of them even invited me to visit, which I thought was just a polite gesture.

A member of that delegation was San-Chiun Shen, who did his PhD with Norman Horowitz at Caltech. He would come to Boston every so often as his son was doing graduate work at the University of Massachusetts. Since his lab also worked on nitrogen fixation, he would take the opportunity to drop in on Fred. Each time, he saw me he extended his but, I had the alibi that I was in the midst of my thesis work. By 1982, however, Fred told him I was near completion, and so he got quite serious about having me teach his lab molecular techniques. What initially was supposed to be a short visit somehow developed into a 1-year plan. About that time, the folks leaving our lab were Sharon Long for a Stanford faculty

[1] A graduate student of that class, Sally Leong, also ended up with a career in USDA.
[2] Also on the same tour was a to-be Harvard classmate Donny Straus.

position, Jonathan Jones to Advance Genetic Sciences, Gary Ruvkun to Wally Gilbert's lab, Venkatesan Sundaresan to Mike Feeling's lab, and Fran DeBruijn to Jeff Schell's empire. Well, you can imagine the response when folks heard of my postdoc in China. Not only did everyone think I was nuts, some even suggested (trying to be helpful) that this might mean an end to my career in the big league. Only Boris Magasanik, a member of my thesis committee, offered supportive advice.

I had to wait for my wife to graduate from Columbia, so I became a postdoc for 6 months in Fred's new lab at the Massachusetts General Hospital. The higher postdoctoral pay was necessary considering my next position in China. I met Stephen Howell at a Keystone Conference and was impressed with his science and personality, so I sent off a bunch of postdoctoral fellowship proposals to join his lab upon my return from Shanghai. The warmest holiday greeting I received in the winter of 1983 was a telegram from the Damon Runyon Foundation. The Helen Hay Whitney Foundation also wanted aninterview but would not pay for my international airfare, so I couldn't go. With a monthly salary of ¥200 RMB (equivalent to ~US $100, but not convertible to foreign currency), it was out of consideration. By early 1984, I also heard from NSF and since it paid more, I declined the Damon Runyon fellowship.

I managed to teach molecular biology techniques through a research project with a graduate student, Yue Xiong, and a lab assistant, Qing Gu. We did a functional analysis of nitrogen fixation gene promoters using site-directed mutagenesis and DNA sequencing. The story behind the story could fill pages; but in short, we completed the work and published in early 1985, surprisingly before a similar paper by a British group later that year. The Editor Rich Losick thought it was the first paper from China in the *Journal of Bacteriology.*

From August 1983 to July 1984, my wife and I lived out of a hotel. With room, board, and roundtrip airfare covered, my monthly ¥200 RMB was just spending money. By Chinese standards, it was a high salary, but we often had to pay tourist prices at hotel stores. Outside of the foreigner-only hotels, many items were rationed, especially food, so money (our nonconvertible type) was worthless without the coupons that were rationed to the Chinese citizens. Of course, while we muse about this, at the heart of the rationing was poverty. What I saw working in the midst of the system was quite a contrast to what I experienced growing up in San Francisco, and I am not from privileged background. As a scientist, you just couldn't help but to question the purpose of science as well as appreciate what food production can mean for others. China today, at least the coastal cities, is much better (largely due to economic reforms). The rural area with approximately half of China's 1.37 billion, however, still has a very long way to go.

After China, I did my NSF postdoc with Stephen Howell at UC San Diego and by the end of 1986, moved to the newly formed USDA Plant Gene Expression Center that is affiliated with UC Berkeley. In 2010, I left the USDA for my current position in Guangzhou, China, the place I set foot in 1978. The scientific career since then has been rewarding and the stories behind them equally interesting. However, the later years were just segments of the journey, on a path that was charted by earlier experiences. Despite living in a publish-or-perish, grant-or-starve environment, I have done my best not to deviate too far from that path, and it is gratifying to know that some of the work bearing my participation have made tangible contributions—the luciferase gene as a research tool, and a transgenic corn product derived from site-specific excision of its antibiotic resistance gene.

Doing well in a career can be less about innate ability, education, and opportunity than with motivation and commitment. Had I not got stuck with the presentation on genetics and agriculture in Hatch Echols' class, I doubt whether this city boy would have taken an interest in agricultural research. Had I not run into Mrs. Bogorad on a sightseeing tour, I might not have spent a year in Shanghai and come away with an experience that has solidified my commitment and priorities in science. Over the years, I had suggested to many graduate students that they ought to consider some postdoctoral time in a less developed country, but few gave it a second thought—aside from thinking that I was nuts.

REFERENCES

Baker SS, Wilhelm KS, Thomashow MF (1994): The 5′ region of *Arabidopsis thaliana* cor15a has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Physiol Biochem* **30**:123–128.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998): Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**:806–811.

Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998): Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression. *Plant J* **16**:433–442.

Goldberg ML (1979): *Sequence Analysis of Drosophila Histone Genes*. PhD thesis, Stanford University, Stanford.

He Y, Michaels SD, Amasino RM (2003): Regulation of flowering time by histone *Arabidopsis*. *Science* **302**:1751–1754.

Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddeloh JA, Riddle NC, Verbsky ML, Richards EJ (2003): *Arabidopsis MET1* cytosine methyltransferase mutants. *Genetics* **163**:1109–1122.

Khoury G, Gruss P (1983): Enhancer elements. *Cell* **33**:313–314.

Kozak M (1986): Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283–292.

Kozak M (1987): At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* **196**:947–950.

Vignali M, Hassan AH, Neely KE, Workman JL (2000): ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* **6**:1899–1910.

Yamaguchi-Shinozaki K, Shinozaki K (1994): A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**:251–264.

CHAPTER 7

Plant Systems Biology

WUSHENG LIU AND C. NEAL STEWART, Jr.

Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

7.0. CHAPTER SUMMARY AND OBJECTIVES

7.0.1. Summary

Plant systems biology attempts to understand the *emergent* properties of plant systems as a whole. Contrary to the hypothesis-driven reductionism in molecular biology, systems biology expands beyond hypothesis-driven science to determine "how plants work" on a grand scale. It is the study of the interactions and dynamic behaviors of the constitutive components, and the establishment of models to monitor and control cellular responses to developmental stages, genetic perturbations, and environmental changes. One focus is to connect "omics" data (genomics, transcriptomics, proteomics, metabolomics) using network analysis and modeling in order to comprehend how biochemicals work in concert in plant biology. The analyses are usually computationally intensive and rely on *bioinformatics* to understand the systems. We will provide definitions and frameworks to understand how various omics data are collected, analyzed, and used in the emerging field of plant systems biology, disciplines, and enabling tools of plant systems biology.

7.0.2. Discussion Questions

1. What is the definition of plant systems biology?
2. How are plant systems biology studies typically conducted, and why?
3. What are plant genomics, transcriptomics, proteomics, and metabolomics?
4. What is bioinformatics? How can bioinformatics be used to help a systems-level understanding of plants?

7.1. INTRODUCTION

Following Francis Crick's discovery of the "central dogma of molecular biology" in 1958 (Crick 1970), biological questions have been mostly studied by a reductionistic approach (Fig. 7.1), which focuses on the functions of individual genes and gene products, one or a few at a time. This approach would be similar to trying to understand how an automobile works by intensely focusing on, say, how the right headlight is constructed. Nonetheless, striving to understanding the function of each gene has

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

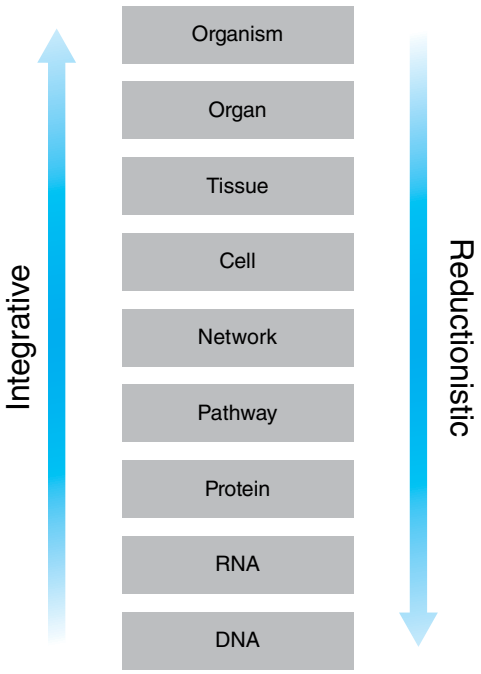


Figure 7.1. Illustration of a traditional reductionistic approach and an integrative approach used in systems biology. (*Source*: Modified from Krepper (2012).)

been greatly beneficial in better understanding the plant as a whole. By understanding gene function, we can better understand biochemical pathways, and vice versa. However, biological systems are well known for their complexity and redundancy. A feature of complex systems is the existence of very large numbers of simple components, such as genes and proteins, which interact extensively to produce complex behaviors. Twenty years ago, the discovery and functional characterization of all the individual genes and gene products in a plant genome seemed like an impossible mission for molecular biologists. However, dozens of plant genomes have now been sequenced using high-throughput methods, and bioinformatics have enabled massively parallel characterization of most, if not all, genes in a genome. The ability to do such large-scale projects marked the beginning of systems biology. For example, the whole genome sequencing of the small plant *Arabidopsis thaliana*, which was the first plant to have its genome sequenced in 2000, revealed that *Arabidopsis* contains greater than 27,000 genes, which encode greater than 35,000 proteins (Arabidopsis Genome Initiative 2000).

Identifying and listing all the genes and proteins that constitute a biological system provides a catalog of the individual components, and is very important for beginning to understand a system's structure. Again, likening to the automobile, it is vital to have the correct parts list to construct a car. However, you also need to know how the parts fit together and wiring diagrams to be able to build the car. Likewise, a plant is not just a combination of genes and proteins, and its properties cannot be fully elucidated from individual components. Understanding the parts and wiring diagrams of plants is the goal of plant systems biology. Thus, an integrative approach to investigate how genes, transcripts, proteins, and metabolite compounds interact with each other in space and time within plants is important for us to understand systems biology (Fig. 7.1).

While scientists have long held an interest in understanding the "big picture" of plant biology, inventions and development in high-throughput biological methods and tools have made systems biology practical as "omics" fields have emerged as some of the most important ones in biology (Fig. 7.2). Genomics, transcriptomics, proteomics, and metabolomics are examples of omics fields

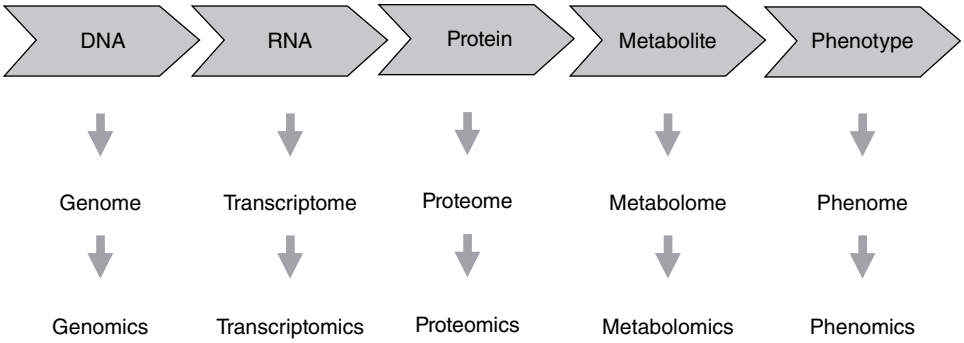


Figure 7.2. The omics sciences in plant systems biology. Genomics focuses on whole genomic DNA and its functional relationships. Transcriptomics investigates messenger RNA (mRNA), which is transcribed from genomic DNA. Proteomics analyzes the expression profiles of proteins (i.e., enzymes and structural proteins), which are translated from mRNA. Metabolomics examines metabolites, which are intermediates and end products of cellular processes. Phenomics studies the phenomes, which are the set of all phenotypes of a cell, tissue, organ, or plant.

that focus on the collection of parts, say, all the genes in an organism (for genomics). The completion of whole genome sequencing and the availability of gene prediction tools pushed all of systems biology ahead and led the way for the other omics because genes are the ultimate source of information to build an organism. The genome-wide analysis of various omics data has generated very large volumes of data sets, requiring rapid progress in bioinformatics and system-level modeling. Thus, the systems view of biology is driven by high-throughput technologies and integrative analysis using computational tools.

7.2. DEFINING PLANT SYSTEMS BIOLOGY

What is a system? As defined by Ludwig von Bertallanfy (1950), a system is a regularly interacting or interdependent group of components forming a unified whole. For example, a house is a system that is composed of different rooms linked by doors or stairs. A plant cell can be viewed as a system consisting of cell wall, plasma membrane, nucleus, chloroplasts, mitochondria, and so on. The most important property of a biological system is the component integration and dynamic interactions among its constitutive components. In addition, a system and the interactions among its components can be distinguished from its environment, with which the system shares input and output relationships.

Then, *what is plant systems biology?* Plant systems biology is the study of the interactions and dynamic behaviors of the constitutive components of a plant system under different conditions, and the establishment of methods and models to monitor and control cellular responses to developmental stages, genetic perturbations, and environmental changes (Fig. 7.3). Thus, plant systems biology is all about networks—how the components of the system (genes, transcripts, proteins, etc.) interact with one another to make a plant phenotype. The main goal of systems biology is to understand biology at the systems level, that is, to obtain a fundamental, comprehensive, and systematic understanding of life (Frazier et al. 2003). To achieve this goal, plant systems biologists need to study: (a) system structures, which include the identity and network of the constitutive components; (b) system dynamics about how a system operates over time under various conditions or perturbations by identifying the underlying mechanism behind the dynamic behaviors of these components; (c) the mechanisms that systematically control the state of the system to minimize malfunctions; and (d) reconstruction strategies to modify and construct biological systems with desired properties (see Chapter 17). Therefore, it is necessary to integrate the information of global DNA, RNA, protein, and metabolite data by combining mathematical modeling and computational analysis with

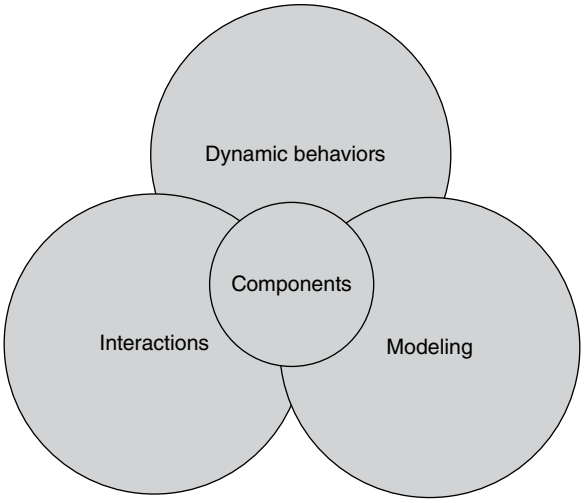


Figure 7.3. Illustration of the definition of plant systems biology. Plant systems biology is the study of the interactions and dynamic behaviors of the constitutive components of a plant system under different conditions, and the establishment of methods and models to monitor and control cellular responses to developmental stages, genetic perturbations, and environmental changes.

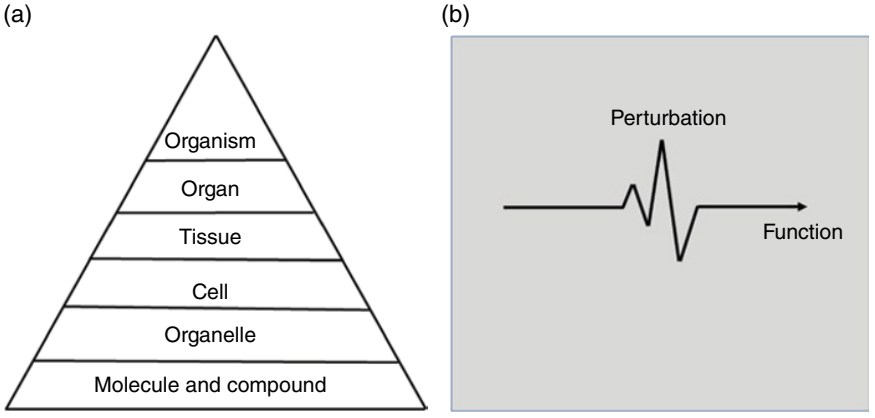


Figure 7.4. Illustrations showing (a) hierarchical properties and (b) robustness that systems biology is used to study. The hierarchy of the various levels one can study ultimately lead to the organism and beyond. Robustness is important to study to understand homeostasis in organisms.

experimentation and hypothesis testing. Once we obtain comprehensive omics data sets, infer the networks of constitutive components, and are able to simulate and monitor the networks, we will be able to understand how these networks give rise to a plant's form and function.

7.3. PROPERTIES OF PLANT SYSTEMS

The study of plant systems is essentially the study of how systems networks are "built" and their dynamic plasticity of response, embodied by three conceptual properties: hierarchy, robustness, and emergence (Lucas et al. 2011).

The concept of hierarchy helps us understand the arrangement of components into different levels within a biological system (Fig. 7.4a). Each level of the hierarchical structure can be viewed as a

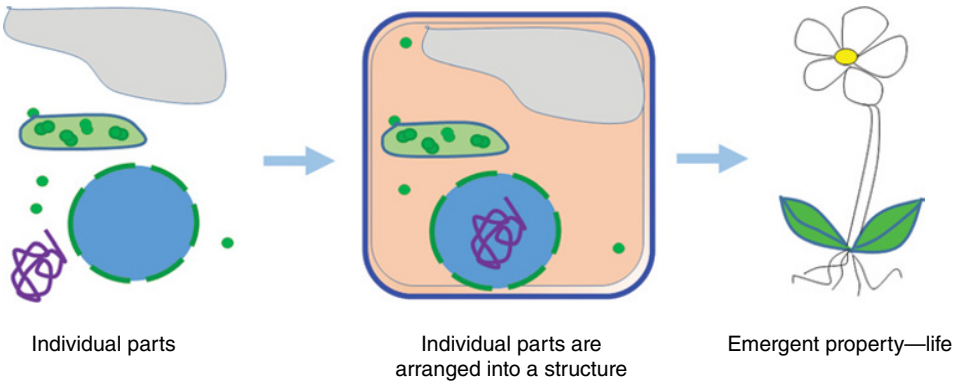


Figure 7.5. Emergent properties arise from the sum of the individual parts. Individual parts are arranged into structure, which determines the behavior and function of the system. Emergent properties are the behavior and function of a system, which depends on properties that go beyond individual components.

system or subsystem. Thus, every system can be viewed as a system of subsystems. For example, plants are made of organs, which consist of tissues; tissues are composed of cells and cells contain different organelles, and then exist levels below organelles. Different hierarchical levels can interact with each other directly (vertically or horizontally) or indirectly. We can choose to study each hierarchical level at a higher or lower level of resolution of analysis, which determines the scale of investigation. The information from each hierarchical level provides new insights into how systems operate as a whole.

The concept of robustness is the capacity of a plant system to maintain functions and behaviors under a wide range of conditions, and the resilience to recover from any changes (Fig. 7.4b). It is an essential property of biological systems and reflects the key characteristics of plant systems: (a) the ability to respond and adapt to environmental changes; (b) the insensitivity to specific kinetic parameters; and (c) the characteristic slow decay process in a system's functions after damage or severe perturbation, rather than a catastrophic failure (Kitano 2002). That is to say, an effective plant systems "knows" when to respond and make changes, and when to keep doing what it was doing before some environmental stimuli ensued. Robustness of a system allows it to maintain function.

Emergence is the novel and irreducible properties at the system level that arise out of the components (Fig. 7.5). Emergent properties cannot be directly achieved by or predicted from the individual components, but instead are reached indirectly by the interactions of the components. They exactly reflect "the whole is more than the sum of its parts." A well-known example is that human consciousness arises from the brain structure, but it is not entirely explained only by brain structure and physiology—at least we do not understand consciousness at these levels. Another example is that plant tissues form into distinct organs such as flowers, leaves, stems, and roots. Each of these plant organs has its own function to play and different characteristics from the individual tissues. Thus, systems biology is especially poised, and likely necessary for us to fully grasp emergence of integration.

7.4. A FRAMEWORK OF PLANT SYSTEMS BIOLOGY

At the very core of the study of plant systems biology is the ultimate goal of being able to understand plant systems sufficiently to enable the construction of predictive models. The focus herein is the structure and dynamics of plant systems. Because of the properties of hierarchy, robustness, and emergence, a systems-level analysis of plants (Fig. 7.6) requires comprehensive data sets for system

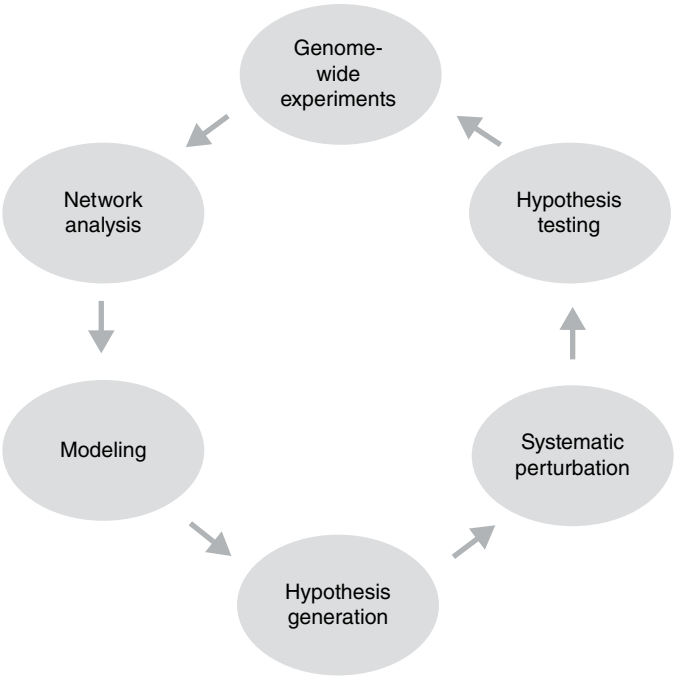


Figure 7.6. The framework of how systems biology is approached. Beginning at the top of the cycle, using multiple high-throughput omics data, network analysis can be inferred and dynamic models can be developed to predict the network structure and behavior. Systematic perturbations can be applied for hypothesis testing and model refinement.

structure analysis, inference of interaction networks among components, and dynamic network modeling, followed by systematic perturbation, model refinement, and hypothesis testing (Ideker et al. 2001; Hood 2003; Palsson 2006).

7.4.1. Comprehensive Quantitative Data Sets

Plant systems biology starts with the acquisition and analysis of very large and, ideally, comprehensive sets of omics data. To obtain a comprehensive data set, plant systems biologists conduct high throughput and accurate measurements to define all of the components of a plant system. Comprehensive plant systems biology requires three types of measurements: (1) a large number of components such as genes, mRNA transcripts, and proteins whose identities and expression level can be measured all at once; (2) time-course measurements for monitoring dynamic changes in time; and (3) spatial patterns of these changes to see where they are occurring in the plant. Systems biology is done on a grand scale.

After large comprehensive data sets are collected using omics methods, computer databases are needed as central repositories for all the data so that other scientists can make use of the information. An important feature of omics data is correct annotation. Annotation refers to assignment of function to the gene, protein, and so on. It is one thing to collect all the A, T, C, G sequence data for genes, but correctly identifying the names and functions of genes is equally important. The availability and accuracy of annotation databases allows scientists to interpret and use information from single genes to complete genomes. For example, if someone clones and sequences a gene, its sequence can be compared to other genes in the Genbank database by a search program called BLAST. Moreover, databases are constantly updated so that database users can automatically access an ever-increasing body of data sets. There are several large databases with genomics, proteomics, and metabolomics data that can be parsed using efficient bioinformatic tools via the Internet.

7.4.2. Network Analysis

The global information of the identity and state of the components is used to predict network structure and dynamics of the systems, which describes the interactions or functional relationships among these components. For example, gene expression and protein synthesis information can be used to predict regulatory relationships between two genes, that is, which one influences expression of the other. The organization of interaction networks underlying a biological system can be represented as mathematical graphs. The simplest way to display such a graph is to use graph nodes to represent a system's components, and graph edges (or links) connecting pairs of nodes to represent the interactions. The nodes of plant system networks can be genes, mRNAs, proteins, metabolites, or other molecules. The edges deduced from high-throughput data are based on correlation analysis.

Gene co-expression networks can be inferred from co-regulated genes that are statistically clustered together: "guilt by association," that is, they respond in a similar manner to certain conditions. Transcriptional regulatory network analysis determines the interactions between transcriptional regulatory genes and their downstream genes. Such signal transduction interactions between transcription factors and promoter *cis*-regulatory elements can be defined using promoter-transcription factor binding assays, computational prediction of *cis*-elements, mutant analysis, and global gene expression profiling. Gene regulatory networks can be studied to reveal how genes regulate other genes' expression patterns by activation or repression. This relationship among genes can be derived from gene expression profiles, mutant analysis, and other data. Networks are often highly dynamic in plants and might be dependent on posttranscriptional regulation such as protein targeting and covalent protein modification. In addition, gene-to-metabolite networks, metabolite-to-metabolite networks, and protein–protein interaction networks can be investigated by focusing on integrative analysis of transcriptome, proteome, and metabolome data.

7.4.3. Dynamic Modeling

The abundance of the components of a plant system can range from a few copies to hundreds or thousands of copies per cell, and always vary in time under different environmental conditions. To capture these dynamic behaviors, interaction networks need to be measured and monitored by quantitative variables indicating the state of expression, concentration, and activity of each node, and by a set of equations indicating the state changes temporally and spatially. Thus, understanding of the networks of component interactions within a plant system requires the use of models. Also, the ultimate goal of systems biology is to be able to model a living organism. Researchers used to generate a biological model to represent a proposed mechanism by sketching a graph using pen and paper. As stated earlier, the simplest graph can represent the system components as a set of nodes that are linked by edges, which indicate the interactions between the nodes. However, because of massive collections of omics data, plant systems biologists must rely on mathematical graph theory to model partial or whole genetic networks.

Integration of comprehensive data sets with prior knowledge can formulate an initial model. The generation of initial models can help to describe the system structure of the interactions and networks that govern the systems behaviors, and to predict relevant dynamics of the system given specified perturbations.

7.4.4. Exploring Systems and Models Toward Refinement

Once an initial systems biology model is developed, specific genetic or environmental perturbations can be applied to the plant system. The corresponding responses to such perturbations can be measured using global discovery tools to capture changes at relevant levels. The detected changes can be integrated with the initial model for model refinement so that its predictions will more closely fit with the experimental observations. When discrepancies arise between predictions and observations,

alternative hypothesis may be proposed, resulting in a refined model. Thereafter, new perturbation experiments can be conducted to further refine the model and distinguish competing model hypotheses.

Thus, a plant systems approach requires that all of the components be investigated globally. Then, the comprehensive omics data are integrated and graphically displayed, and the dynamic networks are modeled to predict the network structure and behavior. Then the model is challenged systematically to understand dynamics and then to refine the model.

7.5. DISCIPLINES AND ENABLING TOOLS OF PLANT SYSTEMS BIOLOGY

Continual advances in technology, sometimes profound, have revolutionized how omics data are collected and analyzed. Advances in molecular genetics and automated nucleotide sequencing have enabled large-scale genetic and physical mapping and whole genome sequencing. Expressed sequence tag (EST) sequencing, microarrays, and RNA-seq have allowed comprehensive mRNA profiling. Advances in mass spectrometry (MS) have enabled the large-scale analysis of cellular proteins and metabolites. Thus, the integrative utilization of all these technologies has advanced the development of plant systems biology.

7.5.1. Plant Genomics

Plant genomics was developed as a discipline that uses genomic DNA sequencing methods and bioinformatics to sequence, assemble, and analyze the structure and function of plant genomes. The term "genome" was created by Hans Winkler in 1920 as a blend of gene and chromosome to represent the complete set of DNA within a single cell. The early days of genomics started with the attempt to shift the focus of molecular biology and genetics from individual genes and proteins to sequencing, mapping, and functional analysis of genomes. Thus, genetics is the hypothesis-driven study of inheritance by linking genes to phenotypes, while genomics is the data- and hypothesis-driven study of genomes which links functions to the information from DNA sequences. Since the term "genomics" was coined by Thomas H. Roderick in 1986, studies on the functions of single genes do not fall into the scope of genomics, unless these studies aim to investigate the effects of single genes on the entire genome's networks.

7.5.1.1. Plant Structural Genomics. Genomics can be divided to be structural genomics and functional genomics, with the former focusing on the mapping and sequencing of genomes and the latter focusing on assigning specific functions to genes. Beginning with the first location of a Duffy blood group gene on human chromosome 1 in 1968, molecular genetics significantly enhanced genome mapping by providing DNA probes for gene identification and *in situ* hybridization, and DNA markers for linkage studies. Mapping of genes on chromosomes can generate genetic and physical maps. A genetic map shows the relative positions of genetic markers along the chromosomes of a species based on recombination frequency, while a physical map is a map of the locations of genetic markers or genes on the chromosomes of an organism, with the distance being measured in base pairs (Fig. 7.7). The highest-resolution physical map would be the complete DNA sequence of a genome that is completely assembled to represent the real DNA configuration in chromosomes.

High-resolution genetic and physical maps can be integrated with shotgun genome sequencing for whole genome sequencing. Whole genome sequencing determines the complete DNA sequence of a genome at a single time point, including the chromosomal, mitochondrial, and chloroplast DNA. Plant structural genomics was revolutionized by completion of the whole genome sequencing of *Arabidopsis thaliana* in 2000, rice (*Oryza sativa* ssp. *Japonica*) in 2002 and poplar (*Populus trichocarpa*) in 2006. The draft genomes of these three model plant species were reconstructed from whole genome shotgun sequencing, sequence assembly, and genetic or physical mapping, followed by gene annotation. Whole genome shotgun sequencing was originally based on the Sanger sequencing method, which was the

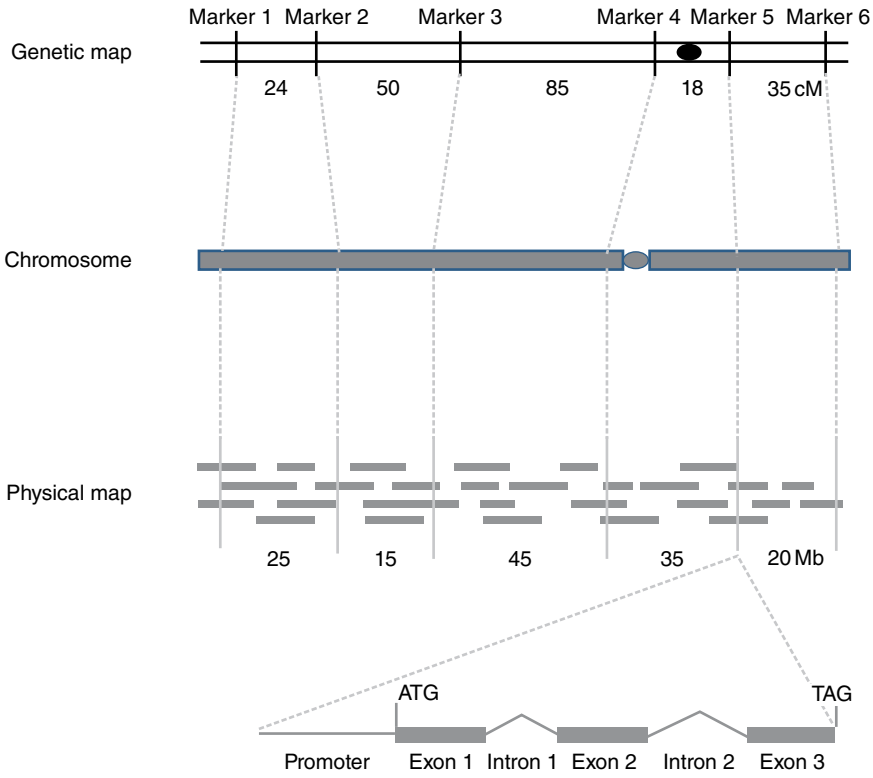


Figure 7.7. Schematic representations of the genetic and physical maps of a hypothetic plant chromosome. The genetic map shows the relative positions of several hypothetical genetic markers along with the relative genetic distances (in cM or centimorgan) among them. The centromere is represented by the circle. The exact physical locations of these genetic markers are indicated in the physical map containing the overlapping clones or the actual genomic DNA sequences (not shown), along with the relative distances between them in millions of base pairs (Mb). The actual distance between two genetic markers may be closer or more distant in the genetic map where there is lower or higher recombination frequency between the markers, respectively. The gene structure of the hypothetical marker 5 is shown.

most advanced DNA sequencing technique until 2005. Then, the first next-generation sequencing (NGS) platforms were commercialized. Sanger sequencing is a DNA sequencing method involving the selective incorporation of radioactively or fluorescently labeled chain-terminating dideoxynucleotides (ddNTPs) by DNA polymerase (Fig. 7.8). Each of these ddNTPs lacks a 3′-OH group required for the formation of a phosphodiester bond with their next nucleotides, causing the termination of DNA extension once a ddNTP is incorporated. Unlike NGS techniques, Sanger sequencing basically reads about 1000 bases at a time. Thus, to sequence a single gene, Sanger sequencing is the way to go. When researchers used Sanger sequencing to read millions of bases, it was very time- and labor-intensive. NGS can read millions of bases simultaneously. NGS typically utilizes Illumina's Genome Analyzer, SOLiD, Roche 454, or PacBio platforms. The invention of new sequencing tools is very dynamic, thus we will not delve into how each machine works to sequence DNA. Suffice to say that the amount of money needed to sequence a plant genome pre-NGS, that is, Sanger sequencing in say 2003, was over 1000 times more than it cost to sequence a plant genome just 12 years later using NGS. Likewise, the bioinformatics used to assemble DNA sequence reads into contiguous pieces of genomic DNA (contigs) are also rapidly changing and is beyond the scope of this chapter to cover. Nonetheless, some bioinformatics is covered later in this chapter.

To conduct whole genome shotgun sequencing (Fig. 7.9), genomic DNA is sheared into random fragments, which are size-selected and cloned into appropriate vectors such as bacterial artificial

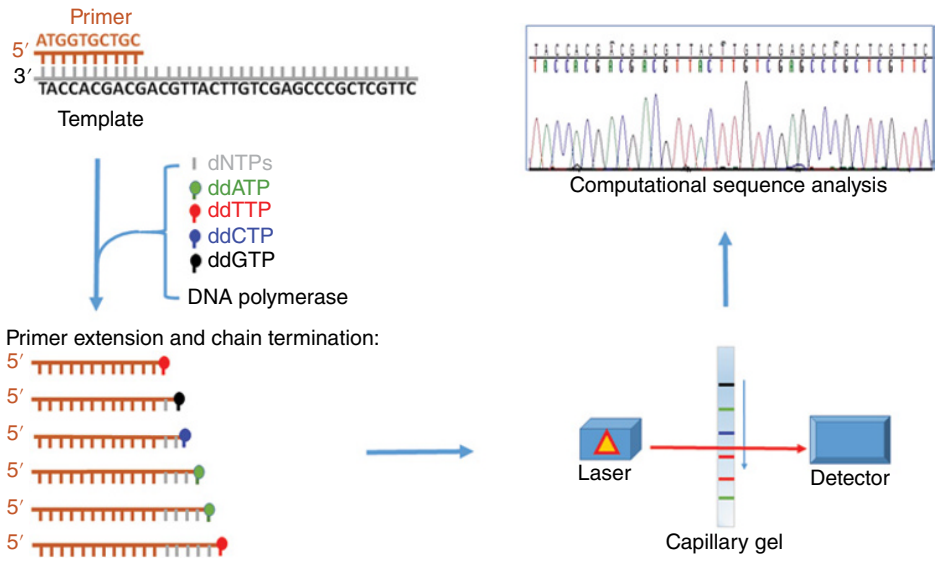


Figure 7.8. The Sanger (chain-termination) method for DNA sequencing. A primer is annealed to a sequence, and reagents are added to the primer and template, including DNA polymerase, dNTPs, and a small amount of all four dideoxynucleotides (ddNTPs) labeled with fluorophores. During primer elongation, the random insertion of a ddNTP instead of a dNTP terminates the synthesis of the chain because DNA polymerase cannot react with the missing hydroxyl. This produces all possible lengths of chains. The PCR products are separated on a single-lane capillary gel, where the resulting bands are read by an imaging system. The sequencing results require computational sequence analysis and storage.

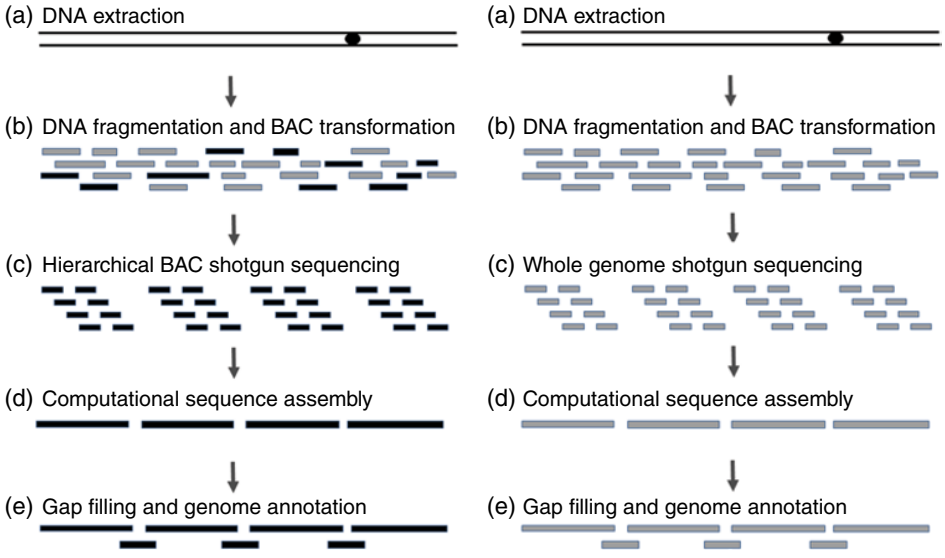


Figure 7.9. Overview of a genome sequencing project showing a hierarchical bacterial artificial chromosome (BAC) method (left) or whole genome shotgun sequencing (right), which contain many of the same steps: (a) Sequencing an entire genome involves the genomic DNA extraction, (b) genomic DNA fragmentation by random sheering and vector construction,(c) shotgun sequencing, (d) computational assembly of the sequences, and (e) gap filling and genome annotation (and verification). The hierarchical BAC shotgun sequencing (left) differs from whole genome shotgun sequencing (right) by selecting a minimal set of overlapping BACs from physical or genetic maps for shotgun sequencing. Putting genomic DNA into BACs is an organizational tool that can be bypassed using the whole genome shotgun approach.

chromosomes (BACs) or a transformation-competent artificial chromosomes (TACs) for sequencing from both ends, resulting in sequence reads. Overlapping reads can be assembled together into "contigs" (i.e., consensus sequences of the overlapping reads) using sequence assembly software, which can then be linked together into scaffolds. Scaffolds can be mapped to genetic or physical maps to determine the sequences in the gaps between scaffolds. With the availability of a draft genome sequence, gene annotation and function prediction can be conducted using ab initio, homology-based or EST-based methods.

7.5.1.2. Plant Functional Genomics. Following the completion of whole genome sequencing of a plant species, the empirical annotation approach shifts its focus on structural genomics to a focus on functional genomics. Plant functional genomics seeks to decipher the functions of unknown genes through an understanding of the relationship between a plant's genome and its phenotypes. It focuses on the properties and functions of the entirety of a plant's genes and gene products by assigning new functions to unknown genes with the help of the vast wealth of data produced by genomic projects. The hypothesized function of an unknown gene can be deduced from algorithms, which allow the comparison of its sequence and presumed protein structure with genes of known functions. But such *in silico* gene function predictions are predictions only—biological experiments are required to determine precise gene and protein function. Close members of a gene family may perform predictable functions, but the exact function of an unknown gene has to be characterized by studying its interactions with other genes in the genome. Thus, a global analysis of temporal and spatial gene expression patterns is needed. Simultaneous analysis of the varying qualitative and quantitative changes of mRNAs (transcriptomics), proteins (proteomics), and metabolites (metabolomics) enables a better understanding of how biological functions arise from the information encoded in a plant genome.

In addition, plant comparative genomics plays an important role in plant functional genomics. The DNA sequences encoding the mRNAs and proteins responsible for conserved functions from the last common ancestors should be conserved in contemporary genomes. It has been found that chromosomal regions in phylogenetically related species may be homologous in terms of gene content. Powerful alignment algorithms are required to align two or more DNA sequences, or even billions of nucleotides between two or more species.

Functional elucidation of genes can also be conducted through systematic perturbations of gene expression such as plant forward and reverse genetic approaches, followed by a whole genome-level analysis of gene expression products such as mRNA, protein, and metabolites. Plant forward genetics uses chemical (e.g., ethyl methanesulfonate (EMS)) or physical (X-rays) mutagens or insertional mutagenesis to randomly generate genomic mutations that could lead to aberrant phenotypes, which, in turn, leads to the identification of the genes that are responsible for the phenotypes after subsequent breeding (Fig. 7.10a). Such mutagenic approaches require time-consuming positional cloning strategies with the help of genetic and physical maps to identify and isolate the target genes, while

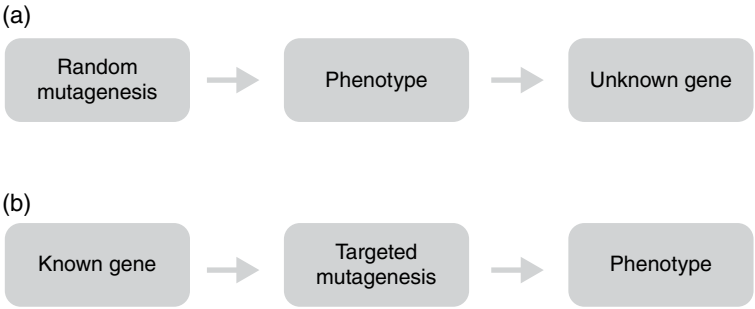


Figure 7.10. Illustration of (a) forward genetics and (b) reverse genetics approaches. Forward genetics uses random mutagenesis to generate mutants and identify the underlying genes whose mutation is responsible for the specific phenotypes of the mutants. Reverse genetics starts with a gene whose DNA sequence is known, and it studies the mutant phenotypes by overexpressing or knocking out the gene.

random insertional mutagenesis allows for a much more direct isolation of the targeted genes with the help of T-DNA or transposons. Plant reverse genetics aims to characterize the exact function of a specific gene with a known DNA sequence by analyzing the phenotypic effects after gene knockout, gene silencing, or overexpression (Fig. 7.10b).

7.5.2. Plant Transcriptomics

Plant transcriptomics is the study of the identities, expression levels, and functions of the transcriptome, that is, all the expressed genes in a tissue or cell type, as represented by RNA. These include mRNAs, rRNAs, tRNAs, and other non-coding RNAs produced in a cell, tissue, or plant during different developmental stages or under various conditions. This term can be used for simultaneous monitoring of the expression profiles of the whole set of transcripts that are actively expressed in a given plant or of a subset of transcripts in a particular cell type or tissue. The plant transcriptome is a dynamic entity that reflects the cell's immediate and genome-wide responses to its environment. Thus, plant transcriptomics aims to (a) catalog all species of transcripts and identify coordinately regulated genes; (b) determine the transcriptional structure of genes, such as transcription initiation sites, 5′ and 3′UTR, splicing patterns, and other posttranscriptional modifications; (c) quantify the dynamic expression levels of each transcript; and (d) dissect the metabolic, biosynthetic, and signaling pathways.

To generate whole-genome expression profiles of all the genes in the cells, various high-throughput technologies are often used, including hybridization- or sequencing-based approaches. Hybridization-based approaches typically hybridize fluorescently labeled cDNA with custom-made microarrays or oligonucleotide microarrays, while sequencing-based approaches determine the cDNA sequences directly using a method called RNA-Seq.

7.5.2.1. DNA Microarrays. DNA microarray technology was derived from Southern blotting, in which a known DNA sequence is used to probe fragmented DNA samples, which are attached to a substrate. A microarray consists of a set of short oligonucleotide probes complementary to the transcripts whose presence and amount is to be investigated, and are immobilized on a solid substrate (Fig. 7.11). Probe design in a microarray is based on known or predicted open reading frames such as ESTs from the target genome. Then, the whole target transcriptome is prepared from samples and converted to cDNAs, which are labeled with fluorescent dyes (either in one color or two), hybridized to the arrays containing the probes, and washed. Because the cDNAs are labeled with fluorescent dyes, light intensity can be scanned and used as a measure of gene expression.

The core principle of microarrays is the hybridization between two DNA strands by forming hydrogen bonds between complementary nucleotide base pairs. A higher number of complementary base pairs in a hybrid sequence means tighter non-covalent bonding between the two strands, thus these hybrid sequences are less likely to be washed off. The strength of the detected signal from a specific spot on the array indicates the presence of the hybrid sequences between the probe and its complementary target, and thus the expression level of that target gene.

Microarray analysis of transcriptomes allows the measurement of low-abundance RNAs with copy numbers of only a few molecules per cell. It is a high throughput and a relatively inexpensive whole transcriptome analysis method. However, microarrays are limited by the available sequence information of the target genome, and may miss exon junctions for novel genes and RNA editing events. Thus, only species with a sequenced transcriptome or genome are appropriate for microarray analysis. And even then, the DNA must be printed onto chips or glass slides. It is difficult for microarrays to detect allele-specific expression. There sometimes exists high background levels owing to cross-hybridization. It is always difficult to compare expression levels among different experiments, which require complicated data processing and normalization. DNA sequence-based methods such as RNA-Seq is rapidly replacing microarrays as the transcriptomics tool of choice.

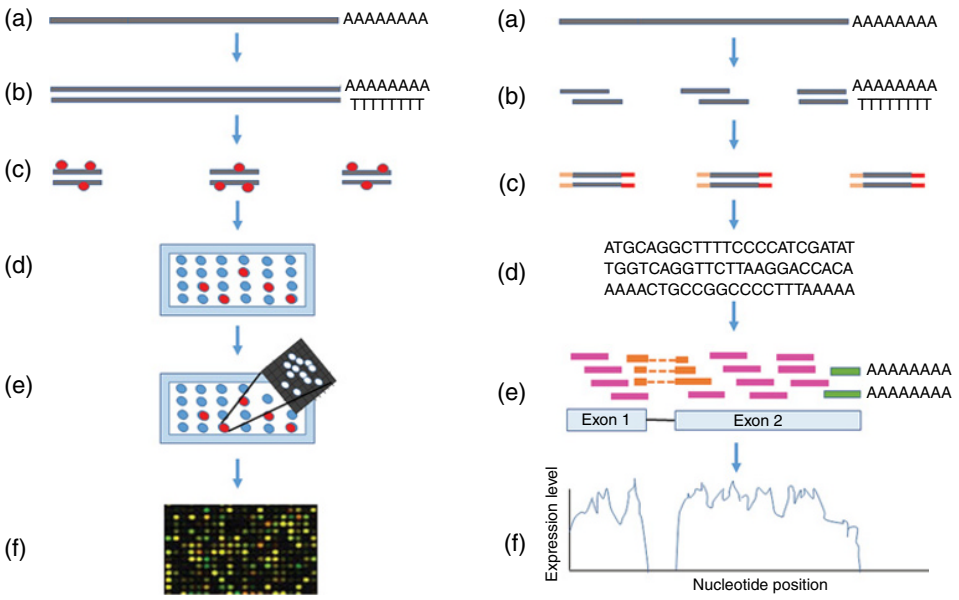


Figure 7.11. Comparison of the flow charts of microarray analysis (left) and RNA-seq (right). To conduct a microarray experiment, the following steps (shown on the left) are taken: (a) total RNA is extracted, (b) which is used for the template for cDNA synthesis, (c) followed by labeling and fragmentation, (d) hybridization and washing, (e) laser scanning, and (f) computer analysis of the expression profiles. RNA-seq shares steps or has analogous steps to microarray analysis, and shown to the right: total RNA is extracted (a) and is fragmented before or after cDNA synthesis (b), followed by ligation to adaptors (c), next-generation sequencing to produce huge amounts of short reads (d). These reads are mapped to a reference genome or transcriptome, or used for de novo assembly, and can be classified as junction reads, exon reads and poly(A) tail reads (e). Then, these reads are used to generate base-resolution expression profiles for different genes (f). (*See insert for color representation of the figure.*)

7.5.2.2. RNA-Seq. A more recently developed deep-sequencing technology called RNA sequencing (RNA-Seq) uses a massive amount of parallel sequencing to conduct whole-genome level transcriptome analysis at a much higher resolution than Sanger sequencing- or microarray-based approaches.

In general, total or fractionated RNA is converted to a cDNA library with adaptors attached to either or both ends (Fig. 7.11). Each cDNA, with or without amplification, is then sequenced in a high-throughput manner to obtain short sequences (typically 30–400 bp in length) from one or both ends using NGS techniques. After transcriptome sequence is obtained, expression profiles can be obtained on a gene-by-gene basis by mapping transcript reads to a reference genome whose transcriptome has been previously studied, or by conducting a *de novo* (out of nothing) transcriptome assembly to produce a genome level transcription map. RNA-Seq has been used to determine transcriptional structure, quantify transcript levels, detect RNA editing events, measure allele-specific expression, and annotate gene structure (5′ and 3′ends, exon/intron boundaries, etc.) without prior genomic or transcriptomic information.

RNA-Seq has become the transcriptome technique of choice, especially to detect unidentified transcripts. RNA-Seq does not depend on hybridization and thus does not suffer from cross-hybridization like microarrays do. It can be very useful to quantify individual transcript isoforms as well as to detect exon/exon junctions. RNA-Seq yields a lot more data than the other methods mentioned earlier and requires bioinformatics tools to extract the desired information about gene expression from the large data sets.

7.5.3. Plant Proteomics

The term "plant proteomics" was first coined in 1997 to represent the study of the identities, contents, structures, and functions of the proteome, that is, the collection of proteins in a plant species, cell, or tissue during various developmental stages and environmental conditions. Proteomes are important when studying biosynthesis pathways and other biological targets, such as specific proteins of interest.

Gene expression, as measured by mRNA quantification, does not always correlate with the synthesis of a specific protein. Protein production from a given amount of mRNA depends on the translation from the mRNA and chemical modifications after translation. Many posttranslational modifications are critical to protein function. These modifications include phosphorylation, methylation, acetylation, glycosylation, oxidation, nitrosylation, and ubiquitination. Moreover, many proteins only function in the presence of other proteins through forming functional complexes. In addition, the rate of protein degradation plays a critical role in the regulation of protein function. Plant proteomics analysis can be conducted to investigate large-scale protein synthesis, structure, posttranslational modification, subcellular localization, and protein–protein interactions. Thus, plant proteomics is an important field in plant functional genomics, and provides a different level of understanding of plant systems.

Each protein has a signature mass, charge, and sequence. Mass and charge can be used in mass spectrometry (MS) techniques to determine which compounds are present in a sample. MS analysis relies on the paths of high-speed gas phase ions of the proteins of interest in electric and magnetic fields, which have their own mass-to-charge ratios and can be distinguished by a mass analyzer. Before passing through the electric and magnetic fields, the compounds have to be vaporized and ionized by either electron impact ionization (EI) or chemical ionization (CI; such as electrospray ionization, atmospheric pressure chemical or photon ionization, and matrix-assisted laser desorption/ionization). Two or more mass analyzers can be used for multiple rounds of mass spectrometry (i.e., tandem mass spectrometry; MS/MS) so that each digested peptide can be broken into smaller fragments, whose spectra can provide effective signatures of the composing amino acids in the peptides leading to protein identification. For example, the first mass analyzer can isolate a compound and determine its mass. The second analyzer stabilizes the gas phase ion of that compound while it collides with a gas and is being fragmented. The last mass analyzer then catalogs the fragments produced from that compound. The charge induced or the current produced by the gas phase ions can be measured by a detector, which produces a mass spectrum for the compound or its fragments.

Using specific performance and mass range of mass spectrometers, two approaches are used to conduct large-scale identification of proteins. The first approach is called the "top-down" strategy of protein analysis, which ionizes the intact individual proteins before being introduced into a mass analyzer. The second approach uses a protease to digest the intact proteins before MS analysis and identification via peptide mass fingerprinting or tandem mass spectrometry. Hence, this "bottom-up" approach uses identification at the peptide level to infer the identity of proteins (Fig. 7.12). One of the "bottom-up" approaches begins with the purification of the cellular proteins, separation through two-dimensional gel electrophoresis (2D gels), excision of individual spots, proteolytic in-gel digestion of the isolated protein in each spot using a protease such as trypsin, and purification of the resulting peptides, followed by peptide MS analysis and computer data analysis against known protein sequence databases. The MS analysis measures the accurate mass of the proteolytic peptides of the isolated proteins by ionizing their chemical compounds in order to generate charged fragments for the measurement of their mass-to-charge ratios (Fig. 7.13). The peptide mass profile of each protein can be compared to that of known proteins in the databases to infer the identification of the target proteins.

Another "bottom-up" approach is shotgun proteomics, which uses a combination of high-performance liquid chromatography (HPLC; Fig. 7.14) with tandem MS (MS/MS). HPLC is a

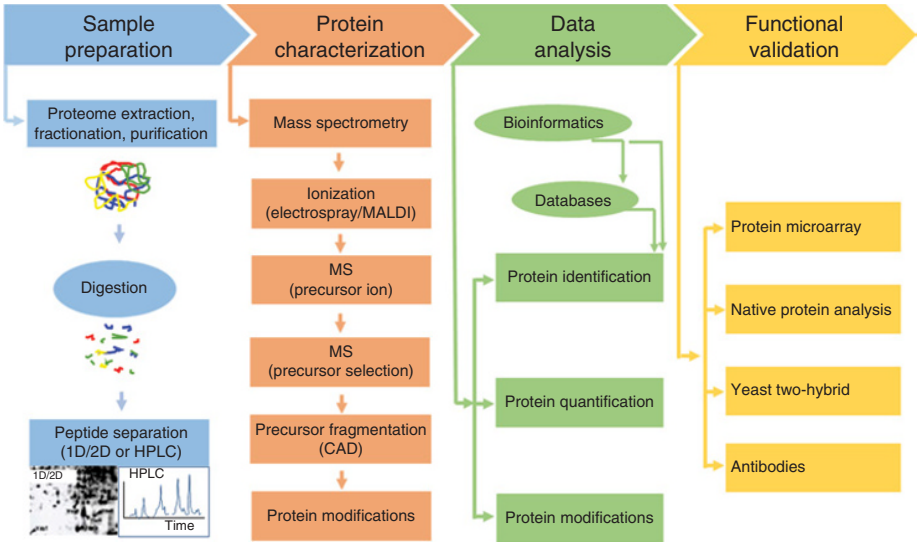


Figure 7.12. An integrated workflow of the bottom-up approach for plant proteomics. This approach consists of steps of sample preparation, protein characterization, data analysis, and functional validation. The samples are prepared by an extraction, fractionation, and a purification step, in which the complexity of the resulting protein samples can be reduced by an electrophoretic or chromatographic separation of protein components, followed by the enzymatic digestion and peptide separation by 1D/2D electrophoresis or HPLC. The resulting individual peptide fractions are subjected to mass spectrometry analysis, which includes ionization of the peptides using (nano)electrospray or matrix-assisted laser desorption/ionization (MALDI), acquisition of a full spectrum, and selection of specific precursor ions based on mass-to-charge ratios, fragmentation through collisionally activated dissociation (CAD), and acquisition of the product-ion spectra (i.e., MS/MS spectra) based on fragment-ion masses. The data are then processed for databases- and/or bioinformatics-assisted protein identification, quantification, and modification analysis. The resulting protein information can be functionally validated using protein microarray, native protein analysis, yeast two-hybrid systems, and antibodies.

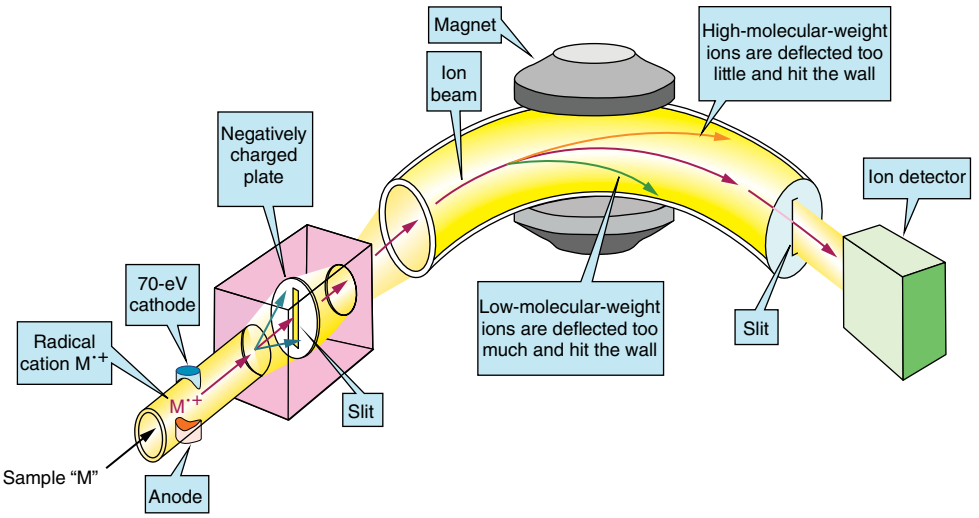


Figure 7.13. Illustration of mass spectrometry (MS). MS instruments are typically comprised three components: an ion source, a mass analyzer, and a detector. The ionizer converts the samples into ions, which are then passed through the mass analyzer, and onto the detector. The mass analyzer sorts the ions according to their mass-to-charge ratio. The detector measures the value of an ion quantity and thus provides data for the abundances of each ion. (*Source*: Reproduced with permission from http://chemistry.umeche.maine.edu/CHY251/Ch13-Overhead4.html)

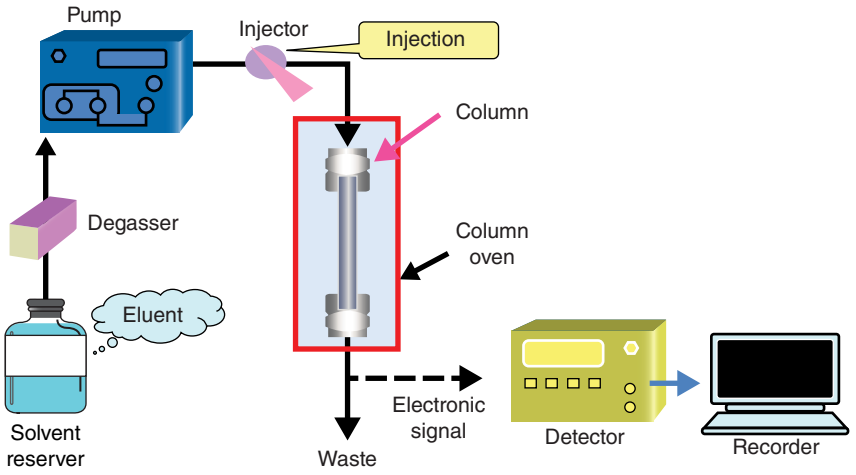


Figure 7.14. Illustration of a high-performance liquid chromatography (HPLC). HPLC instruments typically contain a sampler, pumps, and a detector. The sampler sends the samples into the mobile phase stream. The pumps deliver the desired flow and composition of the mobile phase stream through separation column. The detector generates a signal proportional to the amount of each sample component passing through the column and thus provides a quantitative analysis of the sample components. (*Source*: Reproduced with permission from Showa Denko America, Inc.)

method used to separate, identify, and quantify each component in a mixture. It uses pumps to pass a pressurized liquid solvent containing the resulting peptides after digestion through a microcapillary column filled with a solid adsorbent material. Each peptide interacts slightly differently with the adsorbent material with different hydrophobicity and charges, and flows out the column at different rates, leading to separation of the peptides. Then, they can be further separated by mass-to-charge ratios in two rounds of mass analysis in MS/MS. The mass spectrum of each peptide can be used to identify the source protein from which it is derived by searching against protein databases.

Shotgun proteomics allows global protein identification and quantification, and the ability to systematically profile proteome dynamics. It avoids the poor separation efficiency and mass spectral sensitivity of the 2D gel electrophoresis and MS method, even though shotgun proteomics may maximize the number of identified proteins at the expense of random sampling.

7.5.4. Plant Metabolomics

Metabolomics is the study of the entire suite of metabolites in a plant or tissue. Metabolites are small compounds that are important for plant biological function. It is important to ultimately tie metabolomics to genomics. Metabolites comprise intermediates and end products of cellular processes, including hormones, signaling molecules, and secondary metabolites such as pigments. Primary metabolites refer to those that are directly involved in normal plant growth, development, or reproduction such as amino acids, sugars, vitamins, and so on, while secondary metabolites (i.e., flavonoids, terpenoids, alkaloids, and antibiotics) are produced by modified primary metabolite synthases or from substrates of primary metabolite origin, and are not directly involved in those processes but usually perform important ecological functions. Thus, the plant metabolome is dynamic and context-dependent. It is estimated that *Arabidopsis* has more than 5000 low-molecular-weight compounds; most of them are products of secondary metabolism. Direct chemical analysis of changes in metabolites can be used to link the functional profiles of these small molecules to mutation events.

A global approach may be taken to study as many metabolites as possible at the same time. Otherwise, the first step in conducting a metabolomic analysis is to determine the number of metabolites to be studied. It may be of interest to analyze a defined set of metabolites using a targeted approach, which involves using standard metabolites for comparison. The number and composition of metabolites to be studied determines the metabolic experimental design in terms of sample preparation and choice of instrumentation.

The key to metabolomics is using appropriate chemical profiling technologies, which differentially display or chemically identify the variations in different metabolic states resulting ultimately from genes and gene expression. Such technologies include the development of rapid extraction methods, which should guarantee that the catabolic and anabolic activities of the cells are immediately stopped so that further modifications of the metabolites can be avoided. Metabolites extracted from plant samples are always of great complexity with regard to compound type, which must be significantly reduced before the compounds can be identified and quantified with great precision. Plant metabolites can be separated by using gas chromatography (GC), high-pressure liquid chromatography (HPLC), and capillary electrophoresis (CE), coupled with subsequent MS of the separated molecules. Alternatively, metabolites may be measured directly without chromatographic separation, for example, using nuclear magnetic resonance spectroscopy (NMR) or Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). It should be evident that metabolics depends on many high-end chemistry techniques and a knowledge of chemistry.

To perform GC–MS analysis, samples are converted from solid or liquid phase to a gas phase by being exposed to high temperatures (up to 250°C) before passing through a series of columns. Compounds pass through columns at various speeds (i.e., retention time) because of their specific chemical properties. Each separated compound eluting from the column to the gas phase is ionized before entering the MS so that it can be fragmented in MS by electrons produced by electron impact ionization. The fragmentation patterns are recorded into mass spectral libraries for peak identification. GC–MS is an effective method for chemical analysis because it is relatively easy to use, low in cost, and produces reproducible results with excellent resolutions. However, GC–MS is only suitable for compounds that can be volatilized either at high temperatures or by chemical modifications. Thus, it may not be capable of analyzing some large and polar metabolites.

In another method, HPLC–MS analysis converts the eluting compounds from a solute to a gas phase ion via solute vaporization. It does not require samples to be volatilized so that the compounds are maintained in their native state before separation. It can detect compounds of high molecular mass, great polarity, and low thermostability, which are incompatible with GC–MS analysis. However, HPLC has lower chromatographic resolution than GC. The types of mass spectra produced by HPLC–MS are largely dependent on the instruments used, which limits the construction of reference HPLC–MS spectral libraries because they are often instrument specific.

CE-MS uses gases (i.e., methane or ammonia) to provide the collision energy for fragmentation and the separated compounds do not need to be ionized. It has a higher separation efficiency and sensitivity than HPLC, can measure a wider range of metabolites than GC, and is most suitable for charged analytes. CE-MS produces less reproducible results than GC–MS but is a better approach for quantifying isotopic molecules.

Nuclear magnetic resonance (NMR) is a metabolic fingerprinting method that can characterize metabolites irrespective of size, charge, volatility, or stability. It uses a strong magnetic field together with radio frequency pulses to produce high-energy spin states in the nuclei of the atoms of the samples with odd atomic or mass numbers (e.g., ^{1}H or ^{13}C). The radiation emitted from the nuclei when they return from high-energy spin states to low-energy spin states is detected and used to construct the chemical structure of the compound. NMR allows the samples to be analyzed in intact states and provides high-resolution structural information about the analytes, even though it is less sensitive than other separation methods.

Currently, the major limitation of metabolomics is its inability to profile all of the metabolomes simultaneously due to the chemical diversity, complexity, and dynamic range of the metabolomes. It is also technically challenging to record the steady-state metabolites of plant cells. Current extraction, analytical methods, and instrumentation are inadequate for the identification and quantification of all of the plant metabolites.

7.5.5. Bioinformatics

Following the recent explosive advances in plant omics, and the increased pipeline of omics data, computational tools are needed to handle and analyze the data. Thus, bioinformatics is a quickly growing interdisciplinary scientific field that generates software tools and algorithms to store, retrieve, organize, analyze, interpret, and visualize biological data. It utilizes the many fields of computer science, statistics, mathematics, and engineering to process biological data and generate useful biological knowledge. Thus, the best bioinformaticians are knowledgeable about biology and computer science. The primary goal of bioinformatics is to enable efficient access to, analysis of, and management of various types of information from large data sets, that is, to search and filter massive data sets and acquire informative biological knowledge for logical integration. By doing this, plant bioinformatics enables a systems-level understanding of the functional relationships between genotypes and observed phenotypes as well as building networks as described earlier.

Plant bioinformatics has become a major discipline in almost all the areas in plant systems biology, such as plant genomics, transcriptomics, proteomics, and metabolomics. Advances in whole genome sequencing technologies have provided great opportunities in bioinformatics for the storage, processing, and analysis of the genomic sequences. Bioinformatics develops databases, algorithms, and tools for sequence analysis, such as sequence search and alignment, genome comparison, gene prediction, motif discovery, and phylogeny studies in the field of plant genomics (Table 7.1). Large-scale public databases have been developed and maintained for long-term data storage, such as GenBank (http://www.ncbi.nlm.nih.gov/genbank/) for sequences, UniProt (http://www.uniprot.org/) for proteins, Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) for protein structure information, ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) and Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) for microarray data sets, and Phytozome (http://www.phytozome.net/) for comparative genomic studies in plants. Meanwhile, bioinformatic software have been developed for genomic sequence assembly, comparative genome assembly (i.e., AMOS (http://www.tigr.org/software/AMOS/)), genome annotation, and gene prediction. However, it remains difficult to predict the introns and exons in the DNA sequences, especially for large-sized plant genomes that have experienced rounds of genome duplication. The major limitations in genome assembly and annotation tools remain in the assembly of highly repetitive sequences, which can specifically be addressed using Repeatscout or Smith–Waterman algorithm (Table 7.1). Full-length cDNA or EST sequences or similarities to potential protein homologs in other plant species may significantly increase the precision of gene identification. For example, the GeneSeqer@PlantGDB tool (http://www.plantgdb.org/cgi-bin/GeneSeqer.cgi) has been tailored for applications in plant genome sequencing. In addition, genome comparison tools such as VISTA (http://genome.lbl.gov/vista/index.shtml) can be used to enhance the accuracy of gene prediction. Current limitations of genome annotation and gene prediction include the identification of small genes, noncoding genes, and alternative splicing sites, as well as the accurate prediction of transcription initiation sites.

Sequence comparison allows inference of the function, structure, and evolution of genes and genomes of interest via comparing the sequences, such as a pair-wise, sequence-profile, and profile-profile comparison. BLAST (http://www.ncbi.nlm.nih.gov/blast) and ClustalW (http://bips.u-strasbg.fr/fr/Documentation/ClustalX/) are the most popular tools for pair-wise sequence

TABLE 7.1. Frequently-Used Bioinformatics Tools

Category	Tools	Function	Sources
Database	GenBank	Sequences	http://www.ncbi.nlm.nih.gov/genbank/
	UniProt	Proteins	http://www.uniprot.org/
	Protein Data Bank	Protein structure information	http://www.rcsb.org/pdb/home/home.do
	ArrayExpress	Microarray data sets	https://www.ebi.ac.uk/arrayexpress/
	Gene Expression Omnibus	Microarray data sets	http://www.ncbi.nlm.nih.gov/geo/
	Phytozome	Comparative genomic studies in green plants	http://www.phytozome.net/
Genome sequencing	AMOS	Comparative genome assembly	http://www.tigr.org/software/AMOS/
	VISTA	Comparative genome analysis	http://genome.lbl.gov/vista/index.shtml
	GeneSeqer@PlantGDB	Comparative plant genomics	http://www.plantgdb.org/
Sequence comparison	BLAST	Pair-wise sequence comparison	http://www.ncbi.nlm.nih.gov/blast
	ClustalW	Pair-wise sequence comparison	http://bips.u-strasbg.fr/fr/Documentation/ClustalX/
	PSI-BLAST	Sequence-profile alignment	http://www.ncbi.nlm.nih.gov/BLAST/
	PAUP	Phylogenetic analysis	http://paup.csit.fsu.edu
	PHYLIP	Phylogenetic analysis	http://evolution.genetics.washington.edu/phylip.html
Microarray	AGCC	GeneChip arrays	http://www.affymetrix.com/estore/browse/products.jsp?productId=131429&navMode=34000&navAction=jump&aId=productsNav#1_1
	GeneSpring	Microarray analysis	http://www.genomics.agilent.com/article.jsp?pageId=143&_requestid=1011858
	caARRAY	Web-based data management	http://caarray.nci.nih.gov/
	BASE	Web-based database	http://base.thep.lu.se/
Cis-regulatory element discovery	MDscan	*De novo* motif discovery	http://ai.stanford.edu/~xsliu/MDscan/
	Weeder	*De novo* motif discovery	http://159.149.160.51/modtools/
	MEME	*De novo* motif discovery	http://meme.nbcr.net/meme/
	BioProspector	*De novo* motif discovery	http://ai.stanford.edu/~xsliu/BioProspector/
	Bio-Optimizer	*De novo* motif discovery	http://www.people.fas.harvard.edu/~junliu/BioOptimizer/
	CONSENSUS	*De novo* motif discovery	http://stormo.wustl.edu/consensus/html/Html/main.html
	MotifSampler	*De novo* motif discovery	http://ccmbweb.ccv.brown.edu/gibbs/gibbs.html
	AlignACE	*De novo* motif discovery	http://arep.med.harvard.edu/mrnadata/mrnasoft.html
	W-AlignACE	*De novo* motif discovery	http://www1.spms.ntu.edu.sg/~chenxin/W-AlignACE/
	SCOPE	*De novo* motif discovery	http://genie.dartmouth.edu/SCOPE/

(*continued*)

TABLE 7.1. (continued)

Category	Tools	Function	Sources
Proteomics	SWISS-2DPAGE	2D gel analysis of proteins	http://world-2dpage.expasy.org/swiss-2dpage/
	Melanien	2D gel analysis of proteins	http://au.expasy.org/melanie/
	Flicker	2D gel analysis of proteins	http://open2dprot.sourceforge.net/Flicker/
	PDQuest	2D gel analysis of proteins	http://www.bio-rad.com/en-us/product/pdquest-2-d-analysis-software
	Emowse	MS data analysis	http://emboss.sourceforge.net
	ProteinProspector	MS data analysis	http://prospector.ucsf.edu/
	Mascot	MS data analysis	http://www.matrixscience.com/
	SEQUEST	MS/MS data analysis	http://fields.scripps.edu/sequest/
	Lutefisk	*De novo* sequencing using MS/MS data	http://www.hairyfatguy.com/lutefisk
	PEAKs	*De novo* sequencing using MS/MS data	http://www.bioinfor.com/peaks/features/whatsnew.html
Metabolics	XCMS Online	MS data analysis	https://xcmsonline.scripps.edu/
	MZmine	MS data analysis	http://mzmine.sourceforge.net/
	MetAlign	MS data analysis	http://www.wageningenur.nl/en/show/MetAlign-1.htm
	MathDAMP	MS data analysis	http://mathdamp.iab.keio.ac.jp/
	LCMStats	LCMS data analysis	http://sourceforge.net/projects/lcmstats/
Structure prediction	I-TASSER	Protein structure prediction	http://zhanglab.ccmb.med.umich.edu/I-TASSER/
	HHpred	Homology-based function prediction	http://toolkit.tuebingen.mpg.de/hhpred
	RaptorX	Aligning hard targets	http://raptorx.uchicago.edu/
	MODELLER	Homology model production	http://salilab.org/modeller/
	SWISS-MODEL	Homology model production	http://swissmodel.expasy.org/

comparison. Sequence comparison can be used for building a consensus gene model or homology (relatedness to a common ancestor in evolution) identification. However, pair-wise sequence alignment is insensitive to the detection of distant homologous relationships. Thus, a protein sequence profile can be calculated using the probabilities of the occurrence of each amino acid at each alignment position while multiple protein sequence alignment of a group of closely related proteins is being performed. By doing this, aligned positions in different protein sequences may possibly demonstrate functional and/or structural relationships. PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) is a popular tool for sequence-profile alignment to detect remote homologous relationships. Profile-profile comparison is not widely used to detect remote homologs due to its high false positive rate.

Sequence similarities may not correspond to homology, and may not represent functional conservation. The evolutionary relationships among genes and/or proteins are better represented in a phylogenetic tree. Among the popular methods to generate phylogenetic trees are minimum distance (i.e., neighbor joining), maximum parsimony, and maximum likelihood trees. PAUP (http://paup.csit.fsu.edu) and PHYLIP (http://evolution.genetics.washington.edu/phylip.html) are the most widely used tools for providing options to use any of the three methods for phylogenetic analysis.

Plant bioinformatics also plays a key role in functional analysis, that is, the expression profiling of genes, proteins, and metabolites. Microarray analysis allows the simultaneous measurement of the transcripts of thousands of genes in response to different developmental stages or the environment. The major difficulty of microarray data analysis is the sheer amount of data that should be deposited into a permanent public databases with open access. There are many bioinformatics tools available for a variety of analysis on large microarray data sets, such as Affymetrix's GeneChip Operating Software (GCOS; http://www.affymetrix.com/estore/browse/products.jsp?productId=131429&navMode=34000&navAction=jump&aId=productsNav#1_1), GeneSpring (http://www.agilent.com/chem/genespring), CaARRAY (http://caarray.nci.nih.gov/), and BASE (http://base.thep.lu.se/). Recent efforts in microarray analysis have focused on regulatory sequence analysis, analysis of data across experiments, and correlation analysis of genes that exhibit highly correlated expression patterns under different conditions. Regulatory sequence analysis extracts *cis*-regulatory elements (i.e., motifs) within the promoter regions of co-regulated genes. These motifs are always very short (4-9-necleotide long, typically 6-nucleotide long). There are many different methods for *de novo* motif discovery of statistically over-represented motifs among co-regulated genes such as MDscan Weeder, MEME, BioProspector, Bio-Optimizer, CONSENSUS, MotifSampler, AlignACE, W-AlignACE, and SCOPE (for their applications in plants, see Koschmann et al. 2012; Liu et al. 2014).

Several bioinformatics tools are available to analyze, annotate, and query 2D gel analysis of proteins, such as SWISS-2DPAGE (http://world-2dpage.expasy.org/swiss-2dpage/), Melanien (http://au.expasy.org/melanie/), Flicker (http://open2dprot.sourceforge.net/Flicker/), and PDQuest (http://www.proteomeworks.bio-rad.com). These tools suffer from low accuracy in detecting protein abundance and a limited ability for protein identification. Meanwhile, software for interpreting MS data for protein identification is critical due to the complicacy of the MS data. These include Emowse (http://emboss.sourceforge.net), MS-Fit in the Protein Prospector (http://prospector.ucsf.edu/), and Mascot (http://www.matrixscience.com/). Many tools also exist for MS/MS protein identification (i.e., SEQUEST (http://fields.scripps.edu/sequest/) and Mascot (http://www.matrixscience.com/)), and for peptide *de novo* sequencing using MS/MS data (i.e., Lutefisk (http://www.hairyfatguy.com/lutefisk) and PEAKs (http://bioinformaticssolutions.com/products/peaks)). The limitation in these tools is their incapability to identify proteins because multiple proteins in the databases can fit the detected MS spectra, and to provide the exact sequence of a peptide.

The major challenge in metabolomics is the rapid, consistent, and unambiguous identification of metabolites from complex plant tissues. Computational software is available to identify molecules that vary in samples on the basis of mass and sometimes retention time. Popular metabolomics tools are XCMS Online (https://xcmsonline.scripps.edu/), MZmine (http://mzmine.sourceforge.net/), MetAlign (http://www.wageningenur.nl/en/show/MetAlign-1.htm), and MathDAMP (http://mathdamp.iab.keio.ac.jp/) for mass spectral analysis, and LCMStats (http://sourceforge.net/projects/lcmstats/) for liquid chromatography mass spectrometry. Once the metabolic composition is determined, data reduction techniques such as principle components analysis can be used to elucidate patterns and connections.

Structural biology predicts, simulates, and models the structures of DNA, RNA, and protein as well as their interactions. It also analyzes, infers, and simulates the biological pathways and dynamic networks. Sometimes, a lack of standards for data formats, data processing parameters, and data quality assessment limits accurate, consistent, and reliable data processing and analysis. The leading servers are I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) for protein structure prediction, and HHpred (http://toolkit.tuebingen.mpg.de/hhpred) which is often used for remote homology detection and homology-based function prediction, RaptorX (http://raptorx.uchicago.edu/) for aligning hard targets, and MODELLER (http://salilab.org/modeller/) and SWISS-MODEL (http://swissmodel.expasy.org/) for homology model production.

7.6. CONCLUSIONS

Omics and bioinformatics have made plant systems biology possible. Genome-wide experimental data are generated for the discovery and characterization of all the genes in a genome. Transcriptomics yields information about gene expression, and proteomics and metabolomics data inform about these important compounds that carry out plant biochemical and biological functions. Data collection, analysis, and interpretation lead to the ability to make network models followed by refinement (Fig. 7.1). Since information can be missing because of the presence of false negatives, or misleading because of the occurrence of false positives, caution should be taken to interpret data obtained from each omics approach. Moreover, since data obtained from each omics approach can only provide crude indications of the systematic properties and molecular mechanisms of plant systems, it is critical to integrate data obtained from multiple omics approaches in order to determine the correlations between omics data. The integration of different omics data to plant systems biology has been mainly applied to perform functional annotation of genes, to evaluate putative protein–protein interactions, and to identify components potentially involved in signaling pathways. Such integration has also been applied to examine the topology of networks, providing informative knowledge of the interactions of components, and thus, a systems-level understanding of plant systems.

To take full advantage of plant systems biology, omics technology development and bioinformatics will continue to play critical roles in the post-genomic era. Bioinformatics is important to make biological sense of the massive amount of omics data. The interface of biology and bioinformatics is crucial, and advances in omics and bioinformatics together are needed to make meaningful systems biology models.

LIFE BOX 7.1. C. ROBIN BUELL

C. Robin Buell, Professor of Plant Biology, Michigan State University, Michigan State University Foundation Endowed Professor, Beale Distinguished Faculty.



C. Robin Buell. Courtesy of C. Robin Buell.

When I was growing up, I always wanted to be a veterinarian, as I loved animals. I had no curiosity of plants and my main experience with plants was weeding my family's garden, which clearly was a chore-not a passion. For my undergraduate studies, I attended the University of Maryland as I could not afford to attend an out-of-state school. I majored officially in "Biology" as being "pre-vet" was not really a major at Maryland. There was nothing remarkable about my freshman year, but three major events during my sophomore year changed my whole life, placing me on a career path in the plant sciences. Even today, I am surprised that through a small but seminal set of events as a 19-year old, I abandoned my childhood dream of becoming a vet to beginning a challenging but rewarding career as a plant biologist. I am sure you ask yourself—what could have been that significant? Well, the first event was disappointing experience in an introductory zoology course leading me to question why I wanted to continue with a major in which I had to take another zoology class! This was contradicted by enrollment in the same semester in an introductory botany course with a dynamic, vibrant instructor that made plants exciting! However, even with these two contrasts, it was being hired as a lab assistant in a plant physiology research lab that sealed the deal. In the lab of Dr. Heven Sze, I worked for 4 years washing dishes, assisting graduate students and postdocs, and being mentored by a great group of scientists. Indeed, if it was not for the encouragement and guidance that I had as an undergraduate lab assistant, I doubt I would have been exposed to scientific research and/or careers possible with a graduate degree in plant science.

From this launching pad, I got a position in 1985 as a research technician at the US Department of Agriculture where I was trained in molecular biology, which was in its infancy as a discipline. While I only spent a year as a technician, the skills and knowledge I learned were seminal to my graduate career. I moved west for my graduate work, first getting an M.Sc. in plant pathology from Washington State University and then a PhD in biology/molecular biology from Utah State University. Indeed, a component of my PhD research was sequencing a few kilobases of DNA using Sanger-dideoxy sequencing with 35S, which was state of the art in the early 1990s. During my PhD, I heard about automated DNA sequencing, but it had an error rate of 4% and I thought there was no way that this would ever become established. Well, history has proven me wrong!

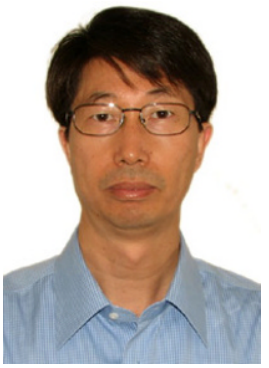
As you know, genomics did not emerge as a discipline until the mid-1990s and as with my undergraduate career path change, becoming a genome biologist was due to a set of unexpected events and opportunities. In 1998, I was an assistant professor at Louisiana State University with a research program focused on *Arabidopsis* molecular plant pathology doing positional cloning when I was invited

to interview for a faculty position at The Institute for Genomic Research (TIGR). At the time, TIGR was one of the premier genomics institutes in the world and a major participant in the Arabidopsis Genome Initiative. Initially, I was not interested in leaving my current position at Louisiana State University but was heavily encouraged by a mentor that this might be a great opportunity. So, I went on the interview and within the first few minutes of being at TIGR, I knew I had to take the position if it was offered. Well, I got the offer and I spent 9 years learning genomics and bioinformatics and doing cutting edge science with a talented group of scientists. As genomics was just developing as a science in the late 1990s, the level of discovery was high. While there, I worked on multiple genome projects enabling the research of a wide number of plant biologists as one key component of all of my work at TIGR was providing large sequence and expression data sets to the public. I feel that the initial funding for plant genomics, in which the funding agencies required immediate public access, was seminal to a majority of the genome-enabled plant science research at the turn of the century. I was fortunate to have been able to play a role in this era.

I have since left TIGR and am now at Michigan State University where I continue my research in plant genome biology and am engaged in a number of productive collaborations with plant biochemists, geneticists, and breeders. I think my life story demonstrates that a career path is not set in stone and that if you seize unexpected opportunities, these cannot only be exciting but also lead you to alternative and rewarding paths in your career.

LIFE BOX 7.2. ZHENBIAO YANG

Zhenbiao Yang, Professor of Plant Cell Biology, University of California, Riverside.



Zhenbiao Yang. Courtesy of Zhenbiao Yang.

I was born in a fishing village on the Southeast coast of China, and grew up during the tumultuous era of the Cultural Revolution, but was fortunate to witness the eve of the Economic Reform orchestrated by Deng Xiaoping during college. I majored in crop pest management from Hainan University (formerly South China College of Tropical Crops), Hainan Island. After finishing my degree there in 1982, I moved to the United States, obtained an MSc degree from Iowa State University in 1986 and PhD from Virginia Tech in 1990, both in plant pathology. After that I did a postdoc in plant signaling at the University of Maryland at College Park, where I discovered Rho GTPases from plants, termed ROPs. After spending 1 year as a research scientist at Virginia Tech, I became an assistant professor at Ohio State University in 1994, where I established pollen tubes and leaf pavement cells as model systems to investigate the ROP-based signaling mechanisms for rapid tip growth and cell–cell coordination of cell morphogenesis, respectively. In 1999, I moved to the University of California at

Riverside and was promoted to full professor in 2003.

My research in pollen tubes uncovered a ROP GTPase-based self-organizing mechanism that controls rapid tip growth. My recent work in the pavement cell system helped to develop a model for a ROP GTPase-centered signaling network of interlinked intracellular interactions and intercellular coordination that underpin the formation of the puzzle piece shape in pavement cells. My research group also helped discover a new auxin signaling mechanism that involves a cell surface auxin perception, and that is distinct from the well-established nuclear auxin perception/signaling system.

My passion for science began early in my childhood by reading popular science books. I became fascinated by nature's mechanisms and the work of scientists to discover the underlying principles of natural occurrences. Some time after that, I read a lot of literature in philosophy and theory and I started thinking more about seeking simple principles behind complex biological systems. My PhD thesis focused on the interaction between potato tubers and the bacterium *Erwinia carotovora*, which causes soft rots in potato tubers and other stored vegetables. I developed a simple system to simultaneously investigate the expression of both bacterial and host genes induced during the interaction. Following the completion of the dissertation, I became interested in cell signaling systems that regulate cellular processes crucial to plant development. My postdoc training focused on the molecular basis of cell polarity in plants, leading to the cloning of Rho GTPase homologs in plants, ROPs (Rho-like GTPases from plants), and the subsequent investigation of signaling mechanisms that regulate the cytoskeleton, cell polarization, polar cell growth, and cell morphogenesis in plants. I currently investigate the coordination of cellular and subcellular activities at the whole tissue and organ level.

During my career, I have published over 90 peer-reviewed research articles and edited 14 books. I have served on numerous departmental and university committees, grant panels, advisory boards, and external committees of research institutions, as well as organized numerous conferences and symposia; have served on numerous editorial boards; and have been an invited speaker at nearly 100 venues.

REFERENCES

Arabidopsis Genome Initiative (2000): Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**:796–815.

Crick F (1970): Central dogma of molecular biology. *Nature* **227**:561–563.

Frazier ME, Johnson GM, Thomassen DG, Oliver CE, Patrinos A (2003): Genomes to life: realizing the potential of the genome revolution. *Science* **300**:290–293.

Hood L (2003): Systems biology: integrating technology, biology, and computation. *Mech Ageing Dev* **124**:9–16.

Ideker T, Galitski T, Hood L (2001): A new approach to decoding life: systems biology. *Annu Rev Genom Human Genet* **2**:343–372.

Kitano H (2002): Systems biology: a brief overview. *Science* **295**:1662–1664.

Koschmann J, Machens F, Becker M, Niemeyer J, Schulze J, Bulow L, Stahl DJ, Hehl R (2012): Integration of bioinformatics and synthetic promoters leads to the discovery of novel elicitor-responsive *cis*-regulatory sequences in *Arabidopsis*. *Plant Physiol* **160**:178–191.

Krepper MA (2012): Systems biology in physiology: the vasopressin signaling network in kidney. *Am J Physiol Cell Physiol* **303**:C1115–C1124.

Liu W, Mazarei M, Peng Y, Fethe MH, Rudis MR, Lin J, Millwood RJ, Arelli PR, Stewart CN Jr (2014): Computational discovery of soybean promoter *cis*-regulatory elements to enable the construction of soybean cyst nematode inducible synthetic promoters. *Plant Biotechnol J* **12**:1015–1026.

Lucas M, Laplaze L, Bennett MJ (2011): Plant systems biology: network matters. *Plant, Cell Environ* **34**:535–553.

Palsson BO (2006): *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, Cambridge, NY.

CHAPTER 8

Recombinant DNA, Vector Design, and Construction

MARK D. CURTIS

Institute of Plant Biology, University of Zurich, Zurich, Switzerland

DAVID G.J. MANN

Dow AgroSciences, Indianapolis, Indiana

8.0. CHAPTER SUMMARY AND OBJECTIVES

8.0.1. Summary

Genomics, biotechnology, and biology in general have been enabled by methodologies to manipulate DNA in a (very tiny) test tube. Restriction enzymes are used as molecular scissors, and ligases are used as molecular "glue." The polymerase chain reaction (PCR) method has become invaluable in amplifying and cloning DNA. In addition, recombination systems have been developed as alternatives to restriction enzymes as rapid and efficient cloning tools. All these methods are useful in creating plasmids containing chimeric DNA constructs that will be transformed into plants.

8.0.2. Discussion Questions

1. What basic elements should be included in the design and construction of an efficient ubiquitous and constitutive plant gene expression vector?
2. Discuss the advantages and disadvantages of recombination cloning technologies versus traditional restriction digestion and ligation technology.
3. Describe a novel strategy to generate a transfer-DNA (T-DNA) vector that allows the expression of several genes from a single position in the genome.
4. Discuss the advantages and disadvantages of using plastid vectors for plant transformation and gene expression.
5. Describe ways in which transgene technology could be made more acceptable to the public.

8.1. DNA MODIFICATION

Recombinant DNA technology relies on the ability to manipulate DNA using nucleic acid-modifying enzymes. The isolation of these enzymes followed shortly after James Watson and Francis Crick's description of the double helical structure of DNA in 1953. Recall that DNA is made up of two

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

twisting complementary strands: alternating units of deoxyribose sugar and phosphates that run in opposite directions. Attached to each deoxyribose sugar is a nitrogen-rich base. The bases, adenine (A), thymine (T), guanine (G), and cytosine (C), on opposite strands are held together by hydrogen bonds to form base pairs (bp); A with T and G with C. The complementary nature of the strands means that each strand provides a template for the synthesis of the other (Fig. 8.1).

In 1955, Arthur Kornberg and colleagues isolated DNA polymerase I, an enzyme capable of using this template to synthesize DNA *in vitro* in the presence of the four bases, in the form of deoxyribonucleoside triphosphates (dNTPs). Although this was the first enzyme to be discovered that had the required polymerase activities, the primary enzyme involved in DNA replication is DNA polymerase III.

While DNA polymerases can replicate a second strand of DNA, they cannot join the ends of DNA together. The discovery of circular DNA molecules (plasmids, discussed later) suggested that such an enzyme must exist. In 1966, Bernard Weiss and Charles Richardson isolated DNA *ligase*, an enzyme that allowed DNA to be "glued" together, catalyzing the formation of a phosphodiester bond (Fig. 8.2).

Soon after this discovery, investigations into bacterial resistance that "restricted" viral growth revealed that *endonucleases* within the cells could destroy invading foreign DNA molecules. Among the first "restriction," endonucleases (referred to as restriction enzymes) to be purified were *Eco*RI from *Escherichia coli*, and *Hin*dIII from *Haemophilus influenzae*. Restriction enzymes went on to become one of the most useful tools available to molecular biologists and deserve special consideration.

Restriction enzymes are produced by a wide variety of prokaryotes. These enzymes are categorized into four different groups (i.e., Type I, Type II, Type III, and Type IV) based on their site of cleavage and the cofactors required for functionality. Type II are the most commonly used restriction enzymes for molecular biology applications. Type II restriction enzymes identify specific nucleotide sequences in DNA of 4–8 bp, usually palindromes, and typically cleave specific phosphodiester bonds in each strand of the DNA within the recognition sequence. The methylation of these specific nucleotide sequences in the host DNA protects the cell from attack by its own restriction enzymes. There are many different site-specific restriction enzymes. These are named after the bacterial species and strain of origin. The restriction endonuclease *Eco*RI, for example, was the first restriction endonuclease identified from the bacterium *E. coli*, strain RY13 (other examples are shown in Table 8.1). Such enzymes recognize a specific double-stranded DNA sequence and cleave the strands to produce either a 5′ overhang, a 3′ overhang, or blunt ends (Fig. 8.3).

DNA fragments that contain single-stranded overhangs ("sticky ends") are the easiest to join together. Two DNA molecules, with compatible single-stranded overhangs, can hybridize to bring the 5′ phosphate and 3′ hydroxyl residues together, allowing DNA ligase to catalyze the formation of phosphodiester bonds (recall Fig. 8.2). In this way, two DNA molecules from different sources can be combined to produce an artificial or "recombinant" DNA molecule (Fig. 8.4). All of biotechnology hinges on recombinant DNA—combining DNA from various sources to do something new. Using two restriction enzymes with different recognition sequences, one can combine two DNA molecules in a predetermined orientation (Fig. 8.5).

The first recombinant DNA molecule was created in Paul Berg's lab in 1972. This pioneering work formed the basis of the recombinant DNA revolution; however, it was not until a year later in 1973 that Stanley Cohen and Herbert Boyer created the first genetically modified organism using these approaches. Combining Cohen's expertise in plasmids and Boyer's expertise in restriction enzymes, a strand of DNA was cut and pasted into a plasmid and maintained and replicated in the bacterium *E. coli*. The transfer of such recombinant DNA molecules to a host cell for amplification is achieved in a process known as *transformation*. Observations in the late 1920s, by Fred Griffith and later by Oswald Avery in the early 1940s, suggested that bacteria could undergo rare natural transformation events. The frequency of these events increased when bacterial cells were treated with cold calcium chloride, which enhanced their competence (i.e., their ability to take up the DNA molecule internally), prior to a brief heat-shock treatment at 42°C. Alternative electroporation



Figure 8.1. The double helical structure of DNA provides the genetic instructions for the development of an organism. The specific sequence arrangement of the bases G, A, T, and C encode regulatory features such as the promoter and terminator sequences of genes, and the triplet code determines the amino acid sequence of proteins. In plants, as with all eukaryotes, most of the DNA is packed into chromosomes and located in the cell nucleus, while in bacteria the DNA is found directly in the cytoplasm and is most often circular.

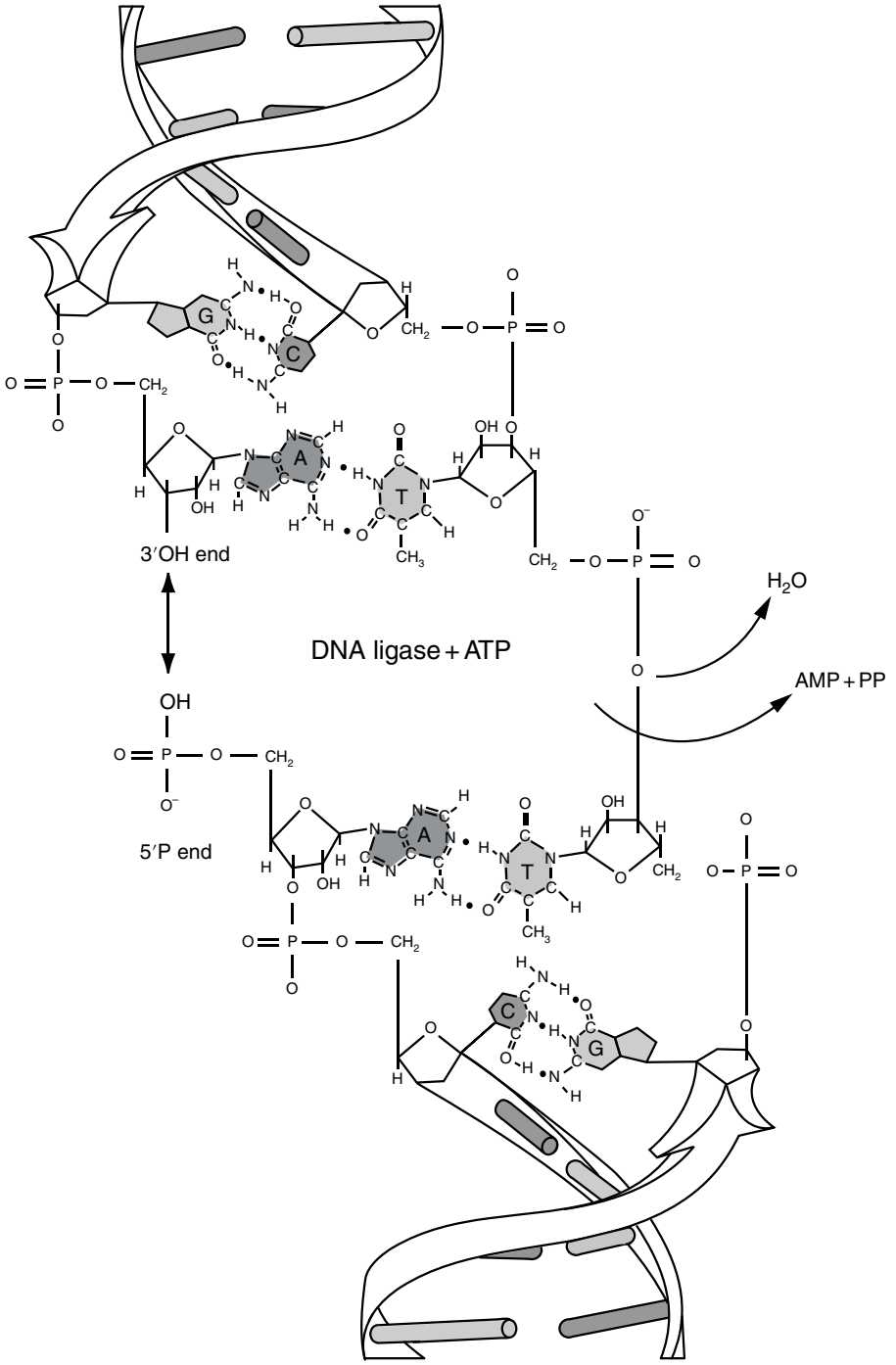


Figure 8.2. The joining of two linear DNA fragments, catalyzed by DNA ligase, creating phosphodiester bonds between the 3′ hydroxyl of one nucleotide and the 5′ phosphate of another.

approaches are now commonly used for transformation. These yield higher transformation frequencies and allow bacterial artificial chromosomes (BACs), too large for conventional transformation, to be taken up successfully by bacterial cells (Sheng et al. 1995). This general procedure formed the basis of clonal propagation, or amplification, of DNA and initiated the development of DNA cloning vectors.

TABLE 8.1. Restriction Endonucleases

Enzyme	Source	Recognition sequence	Cut		Ends
*Eco*RI	*Escherichia coli* **RY13**	GAATTC	G	AATTC	5′overhangs
		CTTAAC	CTTAA	G	
*Bam*HI	*Bacillus amyloliquefaciens* H	GGATCC	G	GATCC	5′overhangs
		CCTAGG	CCTAG	G	
*Hin*dIII	*Haemophilus influenzae* Rd	AAGCTT	A	AGCTT	5′overhangs
		TTCGAA	TTCGA	A	
*Kpn*I	*Klebsiella pneumoniae*	GGTACC	GGTAC	C	3′overhangs
		CCATGG	C	CATGG	
*Not*I	*Nocardia otitidis*	GCGGCCGC	GC	CGCCGG	5′overhangs
		CGCCGGCG	GGCCGC	CG	
*Pst*I	*Providencia stuartii*	CTGCAG	CTGCA	G	3′overhangs
		GACGTC	G	ACGTC	
*Sma*I	*Serratia marcescens*	CCCGGG	CCC	GGG	Blunt ends
		GGGCCC	GGG	CCC	
*Sac*I	*Streptomyces achromogenes*	GAGCTC	GAGCT	C	3′overhangs
		CTCGAG	C	TCGAG	
*Sst*I	*Streptomyces stanford*	GAGCTC	GAGCT	C	3′overhangs
		CTCGAG	C	TCGAG	
*Taq*I	*Thermophilus aquaticus*	TCGA	T	CGA	5′overhangs
		AGCT	AGC	T	
*Xba*I	*Xanthomonas campestris pv. badrii*	TCTAGA	T	CTAGA	5′overhangs
		AGATCT	AGATC	T	

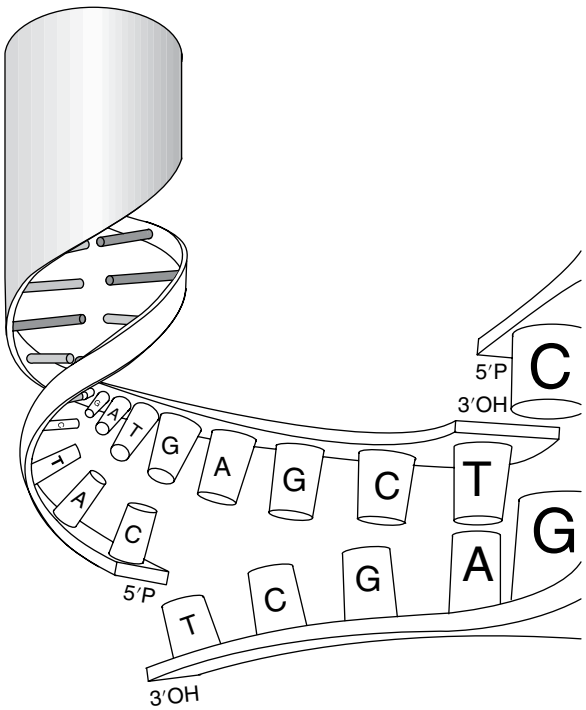


Figure 8.3. The restriction enzyme *Sac*I recognizes a specific 6-nucleotide palindromic sequence wherever it occurs in the DNA and cleaves the DNA asymmetrically at specific phosphodiester bonds to produce 3′ overhangs or "sticky ends."

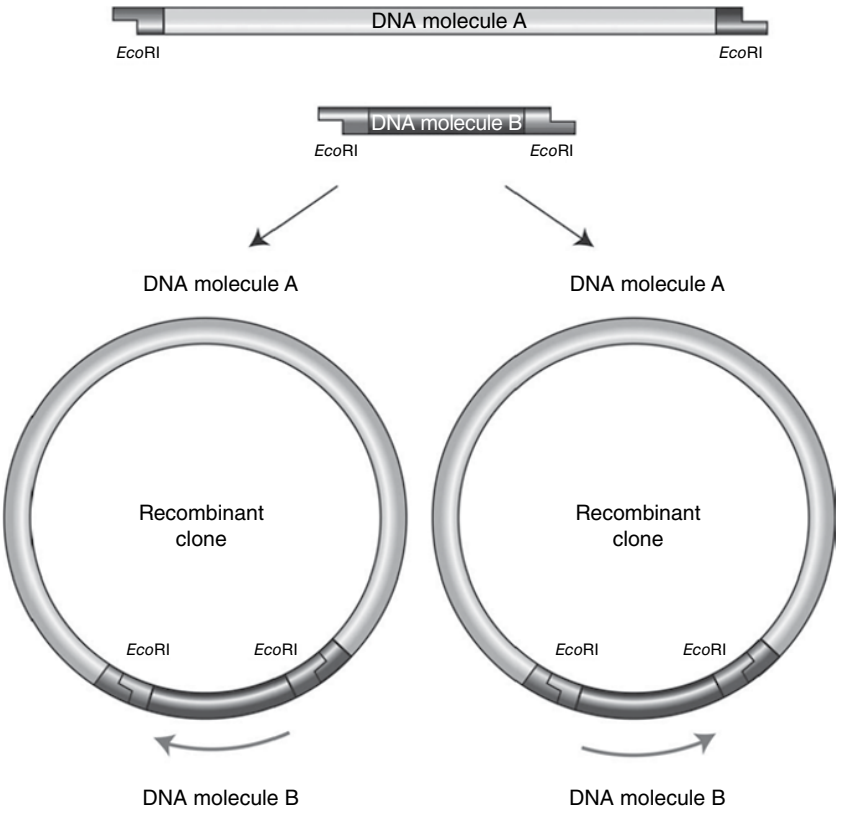


Figure 8.4. DNA fragments produced with a single *Eco*RI restriction enzyme give rise to compatible protruding termini that can anneal in either orientation, bringing together the 5′ phosphate and the 3′ hydroxyl residues on each strand. This allows DNA ligase to catalyze the formation of phosphodiester bonds, joining the two molecules together.

8.2. DNA VECTORS

In molecular biology, a cloning vector is a DNA molecule that carries foreign DNA fragments into a host cell and allows them to replicate. Cloning vectors are frequently derived from *plasmids*, a generic term first coined by Joshua Lederberg in 1952, to describe any extrachromosomal hereditary determinant. Plasmids, found in bacteria but not in plants and other "higher" organisms, are convenient vectors used to manipulate DNA for genetic engineering. Plasmids were discovered in bacteria as double-stranded, extrachromosomal DNA molecules. They have evolved mechanisms to maintain a stable copy number in their host, to ensure that copies are shared between daughter cells and to encode *genes* that provide a selective advantage to their host.

DNA replication determines the plasmid copy number and this is rigorously controlled and closely coordinated with the *cell cycle*. The process of DNA replication is initiated at distinct sites known as *origins of replication* (ori) and proceeds in one or both directions along the DNA. In simple organisms, such as *E. coli*, there is only one origin (oriC); however, more complex organisms, with larger genomes, require many origins to ensure complete DNA synthesis prior to cell division. Origins are usually defined by a segment of DNA, comprising several hundred base pairs, which binds DNA polymerase and other proteins required to initiate DNA synthesis. The plasmid DNA must also replicate in its host organism to ensure that each daughter cell receives a copy of the plasmid. The regulation of this replication determines the number of plasmid copies contained within each cell. Control of plasmid replication is either "relaxed" or "stringent," a characteristic

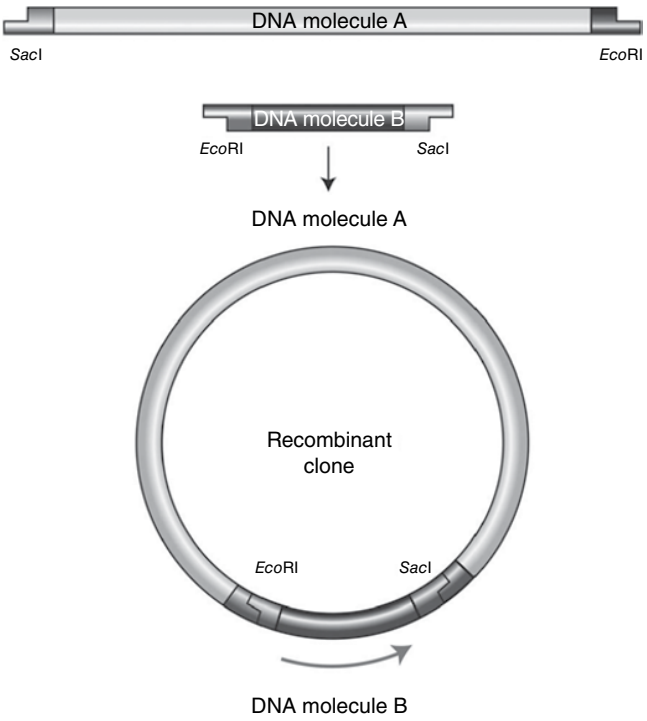


Figure 8.5. DNA fragments produced with two restriction enzymes, *Eco*RI and *Sac*I, give rise to fragments with protruding termini that can anneal in only one orientation with respect to one another, forcing the two molecules to combine in one direction only.

determined by the origin of replication. Plasmids with stringently controlled replication have low copy number, replicating alongside the host's chromosome, once per cell cycle, while plasmids with relaxed replication control have high copy number, replicating throughout the host's cell cycle, resulting in many hundreds of copies per cell. Whether replication is relaxed or stringent, the rate of plasmid DNA synthesis is controlled to maintain harmony with the host's replication. In general, relaxed plasmid replication is controlled by the supply of an RNA molecule, known as *RNA II*, which is required to prime (or start) DNA synthesis (for a review, see Eguchi et al. (1991)). The supply of RNA II is regulated by another RNA molecule, RNA I, which is complementary to the RNA II molecule. When these two molecules hybridize, with the help of a protein known as the *Rop protein*, the priming of DNA synthesis is prevented. Therefore, plasmid replication is inhibited when RNA II is in short supply. Stringently controlled plasmid replication uses a different mechanism. Here plasmid copy number is regulated by the supply of the plasmid-encoded RepA protein, a *cis*-acting protein, which negatively regulates its own transcription and positively regulates the origin of replication (for a review, see Nordstrom (1990)). Relaxed or high-copy-number plasmids are used most often as vectors to produce large quantities of cloned, recombinant DNA, while stringent or low-copy-number plasmids or vectors are used to replicate massive, unstable, foreign DNA fragments such as BACs, or genes that produce lethal effects at high copy number. In addition, low-copy-number *binary vectors* are often used in *Agrobacterium tumefaciens* for transformation of plants to control the number of copies of transfer DNA (T-DNA) which are inserted into the plant genome. This will be discussed in more detail in Section 8.2.1.

Unlike chromosomal DNA, plasmid DNA is dispensable to the host, so why does the host keep it? To be maintained, plasmid DNA molecules must provide their host cells with a selective advantage over their competitors. Plasmid selection is a natural phenomenon that has allowed the evolution of plasmid DNA and its maintenance in bacterial host cells. They encode genes, such as bacteriocins

or antibiotics, enabling the host to kill other organisms competing for nutrients. The first bacterial plasmid identified was the fertility factor (F factor) in *E. coli*, discovered in 1946 by Joshua Lederberg and Edward Tatum. This F factor enables bacteria to donate genes to recipients by conjugation (for a review, see Clark and Warren (1979)), providing a mechanism for adaptive evolution, permitting, for example, plasmid-mediated transfer of antibiotic resistance genes or pathogenicity genes.

8.2.1. DNA Vectors for Plant Transformation

Many bacterial plant pathogens benefit from plasmid-borne pathogenicity genes, which provide them with the ability to infect or parasitize plants. One such organism, *A. tumefaciens*, benefits from a tumor-inducing (Ti) plasmid (Fig. 8.6), which plays a central role in crown gall disease in a wide variety of plants.

The ability of *A. tumefaciens* containing a Ti plasmid to hijack a plant's protein synthesis machinery and genetically engineer the host genome prompted the development of plasmid vectors for *Agrobacterium*-mediated plant transformation. This consisted of removing all of the genes involved in tumor formation and opine biosynthesis within the T-DNA region and replacing them solely with the genes intended for transfer to the nuclear genome of the engineered plant cell (Fig. 8.7). A more detailed description of *Agrobacterium* and elements of the Ti plasmid, as well as *Agrobacterium*-mediated plant transformation, will be discussed in more detail in Chapter 11.

These plant vectors are known as *binary vectors* because they require the interaction of a second, disarmed Ti plasmid lacking a T-DNA. This second plasmid contains the *vir* region, allowing the T-DNA containing the transgenes on the binary vector to be transferred and stably integrated into the host nuclear genome.

Plant binary vectors are constructed and amplified with the aid of *E. coli*, the workhorse organism in molecular biology. Once construction is completed in *E. coli*, such plasmid vectors are transferred

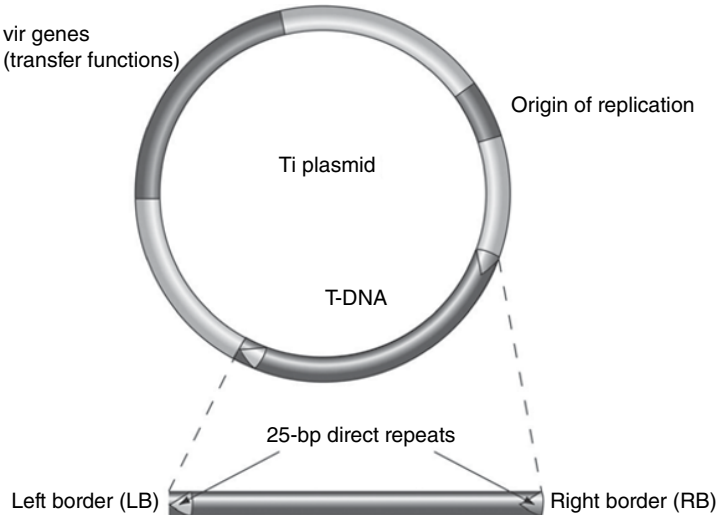


Figure 8.6. The Ti plasmid of *A. tumefaciens* showing the origin of replication, the region encoding the virulence (*vir*) genes, and the transfer-DNA (T-DNA). The T-DNA is flanked by 25-bp direct repeats, known as the *left border* (LB) and *right border* (RB) sequences. The *vir* genes are required for T-DNA processing and transfer to the plant cell. The T-DNA is stably integrated into the nuclear genome of the plant cell, and genes encoded within it, necessary for the biosynthesis of the plant growth hormones, cytokinin and auxin, result in the formation of the characteristic tumorous growth associated with crown gall disease. The T-DNA also encodes opines (nopaline and octapine) that provide the *Agrobacterium* with an exclusive nitrogen source. This provides a competitive advantage to the *Agrobacterium* that carries the Ti plasmid over *Agrobacterium* that does not.

to *A. tumefaciens*, the organism responsible for transferring genes to the nuclear genome of plant cells. These vectors therefore contain origins of replication that function in *A. tumefaciens* and *E. coli*. The pVS1 origin is derived from a *Pseudomonas* plasmid and is stably maintained in a wide variety of proteobacteria, including *Pseudomonas*, *Agrobacterium*, *Rhizobium*, and *Burkholderia*. For this reason, the pVS1 origin has been widely used to construct cloning vectors that are suitable for use in plant-associated bacteria. *A. tumefaciens* uses the repABC operon to stringently control plasmid replication and the partitioning of plasmid DNA to daughter cells. This operon is not only present on large, low-copy number plasmids derived from *Agrobacterium* but also encoded by the chromosomes of *Agrobacterium*. Unfortunately, *E. coli* does not use the repABC operon for plasmid replication, so plasmids containing only the pVS1 origin do not replicate in *E. coli*. Binary vectors designed to shuttle between *E. coli* and *A. tumefaciens* must, therefore, also contain an *E. coli*-compatible *ori*, most commonly the ColE1 origin (providing relaxed replication, Fig. 8.8).

Since plant binary vectors provide no selective advantage to the bacteria, the vectors must be engineered to encode selectable marker genes (e.g., antibiotic resistance genes) for their propagation in *E. coli* and *A. tumefaciens* (examples of commonly used bacterial selectable marker genes are shown in Table 8.2). A broadly active bacterial promoter must be used to transcribe the antibiotic resistance gene, so that bacteria containing the vector can survive and amplify the recombinant DNA. The same selection criteria are used for *E. coli* and *A. tumefaciens*. However, the T-DNA that

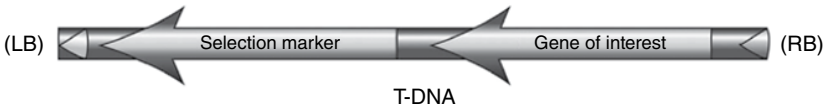


Figure 8.7. T-DNA used to genetically engineer plants frequently contains a selectable marker gene under the transcriptional control of a constitutively and ubiquitously active promoter to ensure gene expression in all tissues at all stages of development, together with the gene of interest (GOI) providing a novel phenotype for the plant.

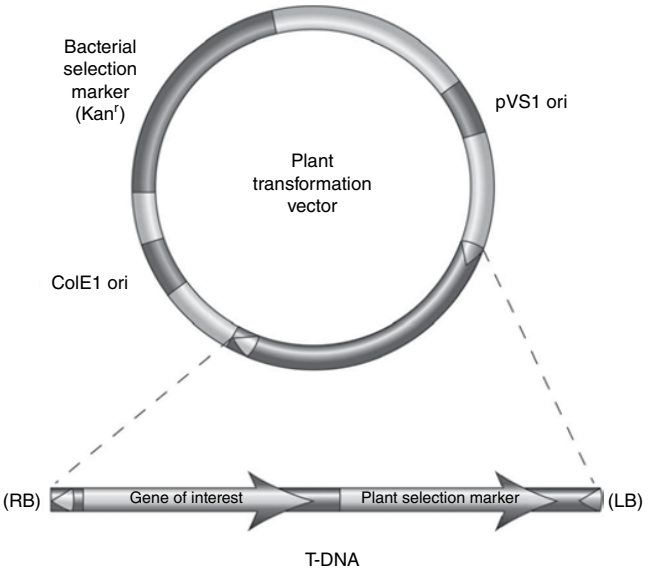


Figure 8.8. A generic plant binary vector with two origins of replication, the pVS1 ori for propagation in *Agrobacterium* and the ColE1 ori for propagation in *E. coli*. The backbone of the vector contains an antibiotic resistance gene for bacterial selection (kanamycin resistance), and the T-DNA contains a plant selectable marker and the gene of interest (GOI).

TABLE 8.2. Commonly Used Bacterial Selectable Marker Genes

Antibiotic	Antibiotic resistance gene	Gene	Source organism
Streptomycin/Spectinomycin	Aminoglycoside adenyl transferase gene	*aadA*	*E. coli*
Kanamycin	Neomycin phosphotransferase gene	*nptII* (*neo*)	*E. coli* Tn5
Chloramphenicol	Chloramphenicol acetyltransferase gene	*cat*	*E. coli* Tn5
Ampicillin	β-Lactamase	*bla*	*E. coli* Tn3
Tetracycline	Tetracycline/H^+antiporter	*tet*	*E. coli* Tn10

is transferred to the plant cell must also contain a selectable marker, this time under the control of a broadly active plant promoter, allowing the identification and propagation of transformed plant cells (Fig. 8.8) (marker genes and the promoters that drive them are discussed in detail in Chapter 10).

8.2.2. Components for Efficient Gene Expression in Plants

The requirements for the successful introduction and efficient expression of foreign genes in plant cells have developed with our understanding of the mechanisms of plant gene expression and plant transformation (for more details, see Chapters 6 and 11). Failure to obtain gene expression using cistrons (gene and promoter sequences) from other species led to the first chimeric genes that used the 5′ and 3′ nopaline synthase (*nos*) regulatory sequences: the nos promoter and nos terminator. Although the nos promoter and terminator sequences are derived from the Ti plasmid of bacterial origin, they share more characteristics with eukaryotic than with prokaryotic genes. The promoter contains sequences that resemble CAAT and TATA boxes, which assist in directing RNA polymerase II (RNAP II) to initiate transcription upstream of the transcriptional start site (Fig. 8.9).

Terminator sequences contain an AATAA polyadenylation signal (which specifies transcript cleavage approximately 30 bp downstream of the signal). Soon after cleavage, multiple adenine residues are added to form a polyA tail on the 3′ end of the transcript. The polyA tail is thought to be important for mRNA stability.

The efficiency of transgene expression in plants is dependent on a number of factors that affect mRNA accumulation and stability. In addition to the promoter (discussed in detail in Chapter 10), these include untranslated sequences (UTRs) both upstream (5′) and downstream (3′) of the gene, codon usage, cryptic splice sites, premature polyadenylation sites, and intron position and sequence (these important features affecting gene expression are discussed in more detail in Chapter 6). Careful consideration of these important factors should be made when designing vectors for transgene expression in plants. Once a decision has been made as to whether a transgene should be expressed ubiquitously or tissue/cell-type specifically, inducibly or constitutively, by changing the promoter fragment used, further decisions can be made that determine whether a gene product is required at high or low levels. Often the omega sequence from the 5′ UTR of the tobacco mosaic virus (TMV) is used to enhance translation in dicot plants. Omega contains a poly(CAA) region, which serves as a binding site for the heat-shock protein, HSP101, which is required for translational enhancement. The efficiency of translation initiation is also affected by other mRNA structures, including the length of the leader—short leader sequences lead to reduced translation efficiency. Secondary structures, both upstream and downstream of the AUG start codon, can inhibit ribosome entry and again reduce translation efficiency. The consensus nucleotide sequence surrounding the AUG start codon in dicotyledonous plants (dicots) is aaA(A/C)aAUGGCu; while in monocotyledonous plants(monocots), it is c(a/c)(A/G)(A/C)cAUGGCG. The presence of upstream AUG codons are particular features of some genes that can reduce translational efficiency (for a review, see Kozak (2005)).

Foreign genes often contain nucleotide sequences that are not commonly used by plants to encode amino acids. Unusual codon usage can affect mRNA stability. For example, *Bacillus thuringiensis* (*Bt*) toxin genes are typically A/T-rich with an A or a T in the third position ("wobble" position) of

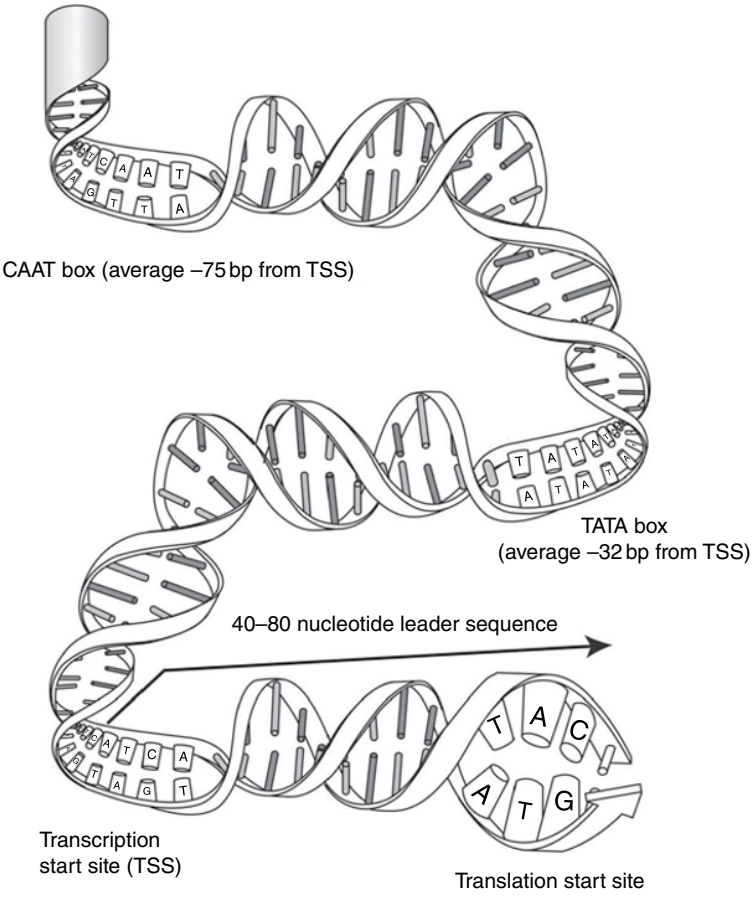


Figure 8.9. Diagram of a generic plant promoter. Typically, transcription factors bind promoter sequences initiating the formation of the transcription complex. Components of the transcription complex bind the CAAT box and the TATA box and assist with the recruitment of RNA polymerase II, allowing the initiation of transcription. The transcription complex can cause the DNA to bend back on itself, bringing together regulatory sequences far from the site of transcription.

codons, which occurs only rarely in plants. Extensive modification of the nucleotide sequence in the coding region of these genes can result in increased expression so that enough Bt toxin would be produced to kill target insects that fed on host plants. The plant species chosen for modification may also influence the design of the transgene construct, since the codon bias in monocot genes tends to be more stringent than it is in dicot genes.

Agrobacterium-mediated plant transformation has had a limited taxonomic host range, with most successful reports of transformation among dicots. Modifications to plant transformation protocols can, however, lead to the successful transfer of genes to plant species once thought to be beyond the natural host range of *Agrobacterium*, including a number of monocots, such as rice, wheat, maize, switchgrass, and many others. Additionally, monocots (and dicots) can be transformed using microparticle bombardment (Biolistics®) (for a more detailed description of microprojectile bombardment-mediated transformation, see Chapter 11). Particle bombardment does not require the use of plant binary vectors containing a T-DNA, since the DNA is physically delivered into the cell by the force of the projected particle. In early plant transformations using particle bombardment, entire plasmids were used, but more recently, only a linear DNA fragment containing the transgene cassette (promoter, gene, and terminator sequences) has been used. This approach has reduced the transgene copy number and eliminated the insertion of unwanted vector backbone sequences.

8.3. GREATER DEMANDS LEAD TO INNOVATION

Recombinant DNA technology has become more sophisticated as new techniques have emerged and greater demands have been made in the analysis of genes and the development of biotechnological innovations. Today it would not be unusual, in the course of analyzing a gene, to express the gene under a variety of promoters, make fusions with reporter genes (Chapter 10) for subcellular localization studies, or make fusions with a purification tag for biochemical analyses. All these types of analysis involve complex DNA manipulations so that a gene and/or its promoter can be inserted into the appropriate vector. Such manipulations have been facilitated by vectors that incorporate a series of restriction endonuclease recognition sites in a sequence known as a *polylinker* or *multiple cloning site* so that there is a convenient place in the vector to insert DNA. However, since vectors do not always contain a standardized polylinker, DNA molecules are not easily exchanged between vector types. In addition, genes and their promoters differ. Genes are rarely flanked by convenient restriction sites for cloning and often contain internal restriction sites that make them incompatible with some vectors. The development of the polymerase chain reaction (PCR) (Fig. 8.10) in 1985, by Kary Mullis, revolutionized the manipulation of DNA, facilitating the inclusion of restriction sites in positions flanking a gene, or its promoter, facilitating cloning as well as the removal of internal restriction sites, while maintaining the integrity of the gene. PCR amplifies specific DNA sequences in a test tube and also allows the sequences to be changed. Despite these improvements, the production of DNA constructs is laborious, and inappropriately positioned restriction sites are still a major factor that hinders vector construction.

8.3.1. "Modern" Cloning Strategies

Several strategies have been developed to overcome the difficulties associated with conventional cloning. These have been compounded by the demands of the numerous functional genomics studies that have resulted from the availability and rapid accumulation of whole-genome sequences. These novel cloning strategies typically fall into two main categories: (1) strategies that rely on *site-specific DNA recombination techniques* and (2) strategies that do not depend on sequence specificity, but instead rely on *DNA end-linking assembly techniques*. Both categories contain a variety of strategies and methods that significantly reduce the time and effort involved in generating recombinant DNA vectors for gene analysis and cDNA library construction (cDNA is a DNA sequence that is complementary to the coding sequence of an RNA transcript). We will discuss in detail the most widely used example of each strategy: Gateway® Cloning (Life Technologies, now part of the company of ThermoFisher) for site-specific recombinational techniques and Gibson Assembly® Cloning (New England Biolabs) for sequence-independent end-linking assembly techniques. Additional cloning methods worth mentioning include Golden Gate and BioBricks that utilize Type IIs restriction enzymes for cloning. Because these unique restriction enzymes cut outside of their recognition sequences, they allow for simultaneous ligation of multiple DNA fragments within the same reaction; and in the process, they have the added benefit of removing the restriction enzyme recognition sites from the final DNA molecule. For more information on the current status of cloning methods, see the references mentioned below or recent reviews by Ellis et al. (2011) and Patron (2014).

8.3.1.1. Gateway Cloning. Many site-specific recombinational cloning methods have been developed in the past two decades including Cre-recombinase-based systems such as Creator™ and Echo™ cloning (previously available from Clontech and Life Technologies, respectively), or more recent systems such as Gateway (Life Technologies), Golden Gate (Engler et al. 2008) and BioBricks cloning (Shetty et al. 2008). Advantages vary depending on each individual system, but these methods all offer accurate and efficient recombinational cloning of multiple DNA fragments within one reaction based on specific recognition sites within the sequence. In the case of Creator, Echo and Gateway cloning, this reaction does not rely on restriction enzymes or ligase, but instead uses a

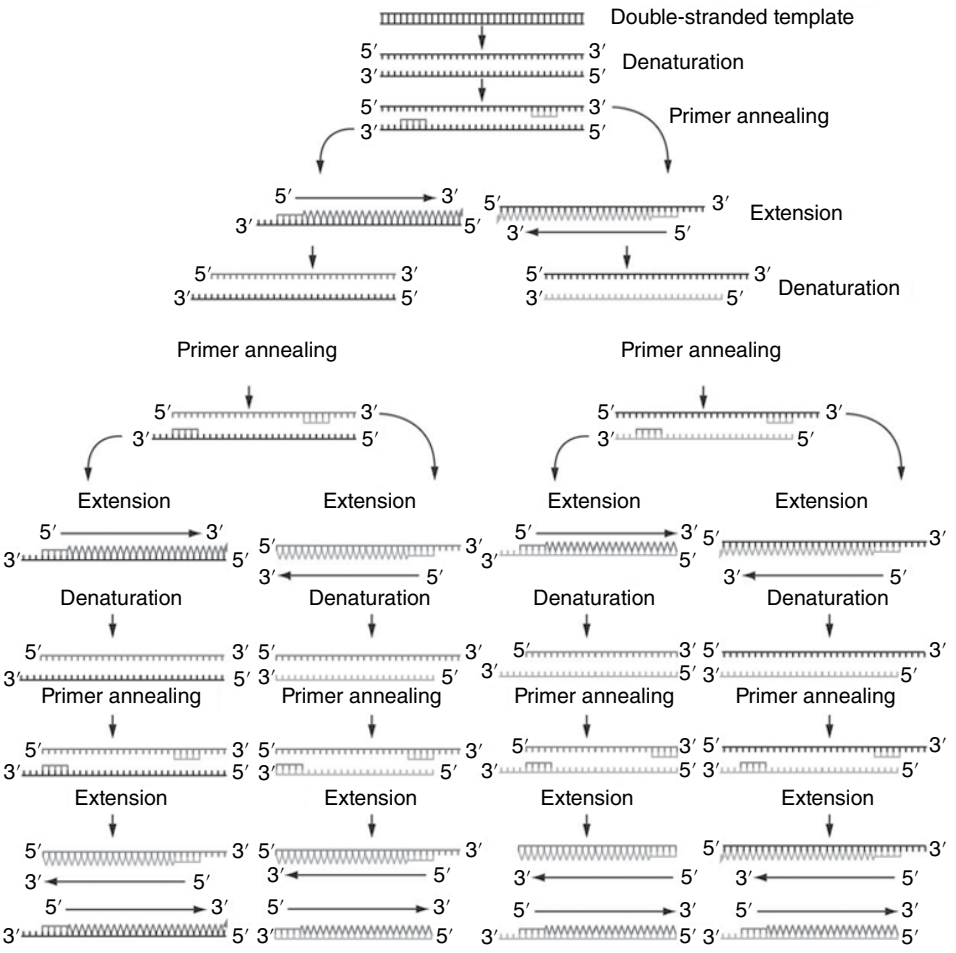


Figure 8.10. Polymerase chain reaction (PCR) is a technique that allows a chosen region of DNA to be amplified *in vitro* by separating the double-stranded DNA template into two strands by denaturation and incubation with oligonucleotide primers and DNA polymerase to synthesize a complementary strand of each. The primers can be designed to incorporate restriction enzyme recognition sites or any other recognition sequence to facilitate the cloning of PCR fragments. Repeated cycles of denaturation (heating), primer annealing (cooling), and extension (heating) for DNA synthesis with DNA polymerase allow the targeted region of DNA to be amplified many thousands of times. This tool is frequently used in biotechnology, forensics, medicine, and genetic research to amplify DNA fragments.

recombinase enzyme (i.e., Cre recombinase or Integrase) for recognition of a unique site sequence within the DNA and subsequent recombination to produce the desired DNA molecule. To gain an in-depth understanding of how site-specific recombinational cloning works, the Gateway cloning system will be discussed in more detail in the following. The Gateway cloning system takes advantage of elements that evolved naturally in the life cycle of the bacteriophage *lambda* (λ). During this cycle, the bacteriophage passes from a lysogenic phase, in which the viral genome is stably incorporated into the host genome, to a lytic phase, in which the host cell ruptures (lyses) and infectious phage particles are released (for a more recent review of lambda development, see Oppenheim et al. (2005)) (see Fig. 8.11).

The Gateway cloning system utilizes modified *att* recombination sites, together with an integration enzyme mix containing Integrase (Int) and Integration Host Factor (IHF) proteins (BP clonase) and an excision/integration enzyme mix containing the Int, IHF, and Excisionase (Xis)

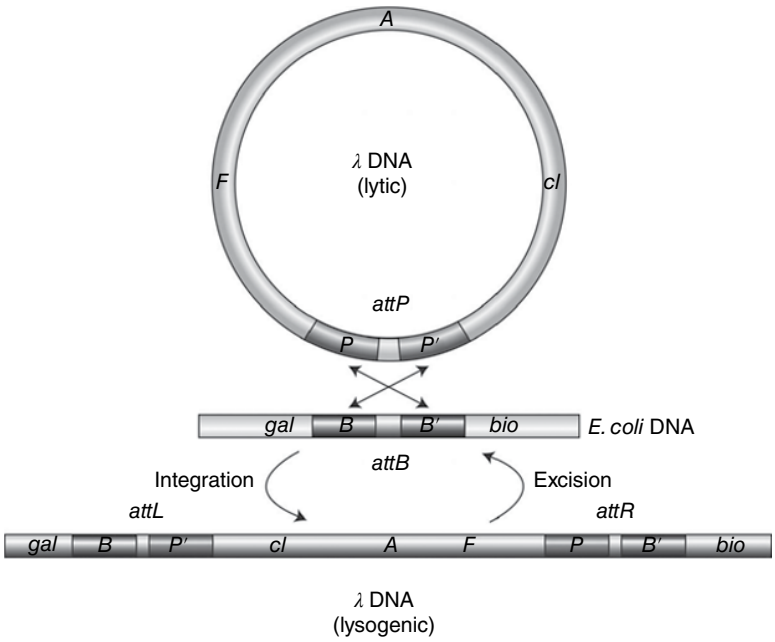


Figure 8.11. For lysogeny, the viral DNA is incorporated into the host genome by a process of recombination between common sequences, the *att* sites, in the two genomes. Bacteriophage λ contains an *attP* site (*P* for phage), and the host *E. coli* DNA contains an *attB* site (*B* for bacterium). A number of proteins are required for this recombination: λ-derived integrase (Int) and *E. coli*–derived integration host factor (IHF) allow λ to enter the lysogenic phase of its life cycle, and IHF, Int, and λ-derived excisionase (Xis) allow λ to excise from the *E. coli* genome and enter the lytic phase of its life cycle.

proteins (LR clonase), derived from elements used during the bacteriophage λ life cycle. DNA fragments flanked by recombination sites can be mixed *in vitro* with vectors that also contain recombination sites, allowing the exchange of DNA fragments and the generation of recombinant DNA. Such an approach avoids many of the difficulties associated with conventional cloning (inconvenient restriction sites, time-consuming reactions, etc.). For Gateway cloning, the *att* sites have been modified so that the orientation of the DNA fragments can be maintained during the excision and integration process. Catalyzed by BP clonase, an *attB1* site specifically recombines with an *attP1* site to produce *attL1* and *attR1* sites, while an *attB2* site specifically recombines with an *attP2* site to produce *attL2* and *attR2* sites (Fig. 8.12). This allows PCR fragments flanked by *attB1* and *attB2* sites to be inserted into pDONR vectors containing the reciprocal *attP* sites, thereby generating "entry clones" in which the chosen DNA fragments are flanked by *attL1* and *attL2* sites.

Entry clones should be sequence-validated to provide a library of well-characterized DNA fragments for insertion into "destination vectors." Catalyzed by LR clonase, DNA fragments flanked by *attL1* and *attL2* sites are then transferred, by a second recombination reaction, to pDEST vectors containing *attR1* and *attR2* sites. The resulting recombinant DNA constructs are known as "expression clones." Here, the recombination product of the *attL1* and *attR1* sites are *attB1* and *attP1* sites and the recombination product of the *attL2* and *attR2* sites are *attB2* and *attP2* sites (Fig. 8.13).

To select the correct recombination product for the BP and LR reactions, a combination of positive and negative selectable markers are employed. Positive selection is afforded by alternative antibiotic selection, while negative selection is afforded by the *ccdB* gene, the product of which inhibits the activity of DNA gyrase, thus preventing negative supercoiling during DNA replication, and ultimately leading to cell death. *E. coli* bacteria transformed with vectors containing the *ccdB*

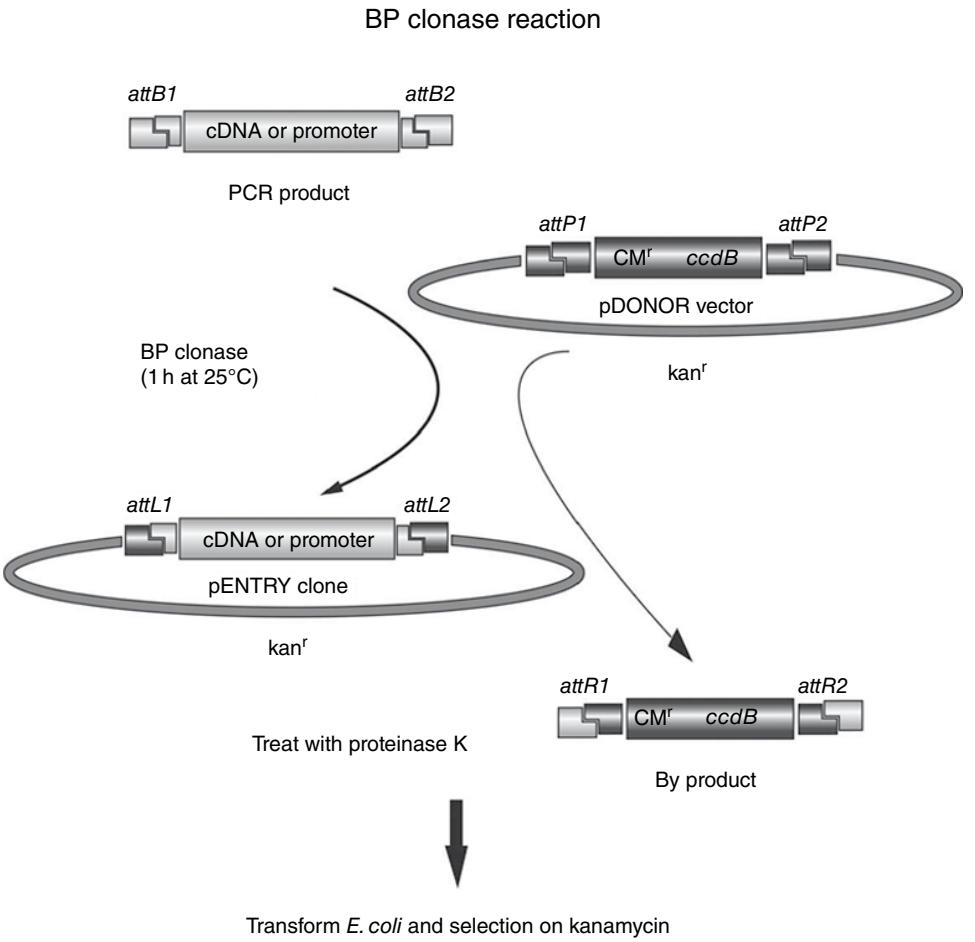


Figure 8.12. A gene or promoter is amplified by PCR using DNA target-specific primers that contain the *attB* sites (*attB1* and *attB2*) at the 5′ and 3′ ends, respectively. The purified PCR product, flanked by *attB* sites, is mixed with a pDONOR vector that contains the corresponding *attP* sites. To this DNA mix is added BP clonase enzyme (containing Int and IHF proteins). After 1 h incubation at 25°C, proteinase K is added and incubated for a further 20 min at 37°C. This mix is used to transform *E. coli* bacteria and plated on the appropriate antibiotic (in this example, kanamycin) selecting transformants containing the appropriate pENTRY clone.

gene (i.e., pDONR or pDEST vectors) or by cointegrate intermediates cannot grow. Only bacteria containing the desired recombinant construct that lacks the *ccdB* gene and contains the appropriate antibiotic resistance marker gene can survive (Figs. 8.12 and 8.13). The propagation of pDONR vectors and pDEST vectors is achieved using the *E. coli* strain DB3.1, which contains a mutant DNA gyrase, which is unaffected by the *ccdB* gene product.

8.3.1.2. Gibson Assembly Cloning. DNA end-linking assembly methods have many advantages over traditional restriction digest and ligation cloning: (a) no restriction enzyme digest is required, (b) no compatibility of the DNA fragment ends is needed (as shown in Figs. 8.4 and 8.5), (c) multiple DNA fragments can be efficiently assembled together simultaneously, and (d) there are no remaining restriction enzyme site "scars" or "seams" remaining after assembly (hence the general method has been referred to as "scarless" or "seamless" cloning. Many variations of this DNA assembly method exist (Gibson Assembly (New England BioLabs), In-Fusion® (Clontech),

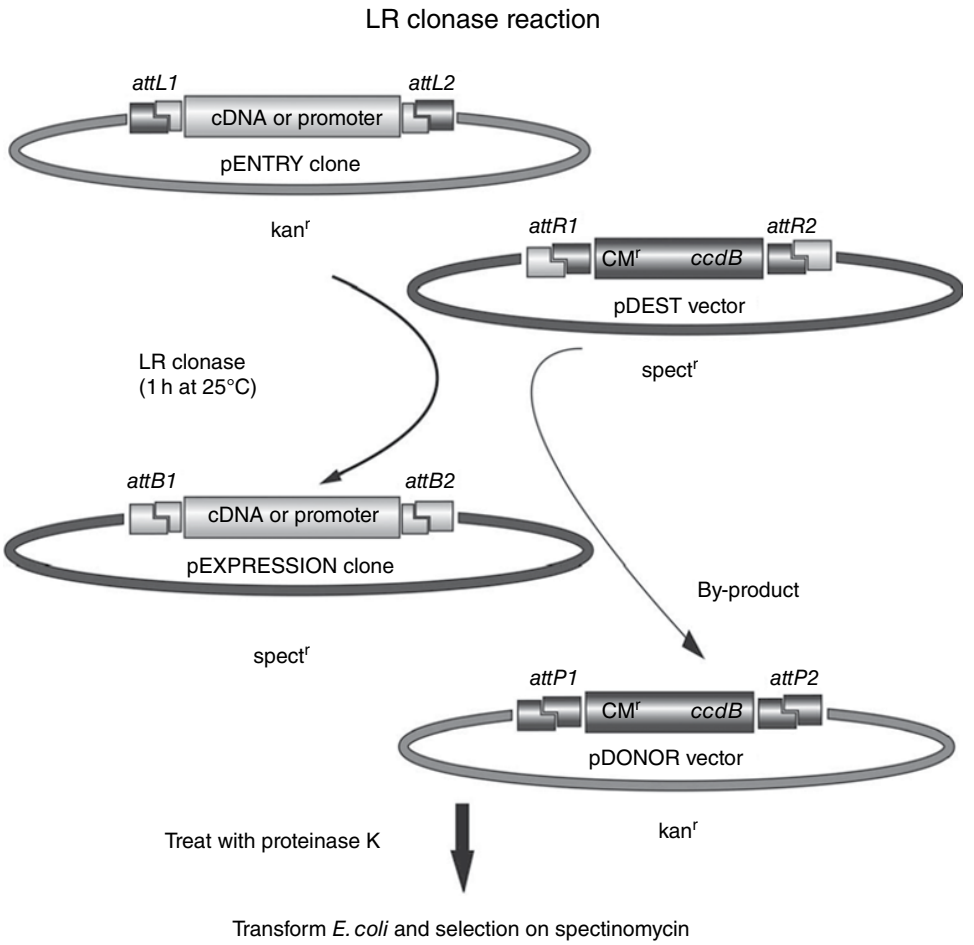


Figure 8.13. A gene or promoter contained within the pENTRY clone flanked by *attL* sites (*attL1* and *attL2*) is mixed with a pDESTINATION vector that contains the corresponding *attR* sites. To this DNA mix is added LR clonase enzyme (containing the Int, IHF, and Xis proteins). After 1 h incubation at 25°C, proteinase K is added and incubated for a further 20 min at 37°C. This mix is used to transform *E. coli* bacteria and plated on the appropriate antibiotic (in this example, spectinomycin) selecting transformants containing the appropriate pEXPRESSION clone.

GeneArt® Seamless cloning, SLIC method (Li and Elledge 2007)). However, Gibson Assembly cloning is the example that we will discuss here in more detail.

Gibson Assembly cloning consists of an exonuclease, a high-fidelity proofreading DNA polymerase, *Taq* DNA ligase, and multiple overlapping linear DNA fragments. In this reaction, linear DNA fragments with overlapping ends can be created through PCR amplification, restriction enzyme digestion, or *de novo* synthesis (i.e., generation of DNA which does not require a template sequence). The overlapping ends of DNA are crucial for this cloning method, as they provide the homologous regions that will be recognized by the enzymes and allow for the "stitching" together of linear fragments into a seamless DNA molecule. Once the overlapping linear DNA fragments are ready, they are joined with the exonuclease, polymerase, and ligase enzymes into one reaction mixture. Within this reaction, the exonuclease functions by cleaving nucleotides from the ends of the linear DNA, leaving an overlap of homologous sequences in the double-stranded fragments and allowing these

fragments to anneal together. Subsequently, the high-fidelity proofreading DNA polymerase fills in any remaining gaps in the annealed DNA fragments (the proofreading nature of this polymerase ensures a low rate of errors in the sequence), and the *Taq* DNA ligase "glues" the final molecule together (as shown in Fig. 8.2). Final reactions can simply be transformed into *E. coli* and plate on the appropriate antibiotics to select for positive colonies.

This method was originally described by Daniel Gibson and his colleagues while assembling large circular DNA molecules (i.e., the 583 kb genome of *Mycoplasma genitalium*) at the J. Craig Venter Institute (Gibson et al. 2008). Since originally reported, the method has been improved upon by using a 5′ T5 exonuclease III and combining the reaction into one single isothermal step (Gibson et al. 2009), as well as reducing the overlapping DNA regions from 450 base pairs to 15–20 base pairs. In this improved version, the entire reaction is carried out at 50°C. At this elevated temperature, the 5′ T5 exonuclease is able to cleave the linear DNA fragments from the 5′ to the 3′ end (which has the added benefit of not interfering with DNA polymerase) and is eventually inactivated within the reaction as a result of its heat-sensitive nature at 50°C. Additionally, the T5 exonuclease is limited to cleavage of linear fragments, but cannot cleave circular fragments. These characteristics allow the exonuclease enzyme to be combined within the same reaction with the DNA polymerase and *Taq* DNA ligase, which both perform their enzymatic functions efficiently at elevated temperatures (e.g., 50°C). While the Gibson Assembly cloning method has been used for the simultaneous end-linkage and assembly of multiple DNA fragments in one reaction (e.g., 38 different 60 base pair sequences were correctly assembled and validated in yeast cells in one reaction (Gibson 2009)), extremely large DNA molecules have required multiple assembly stages (the assembly of the *M. genitalium* required five successive stages).

8.4. VECTOR DESIGN

Recombinant DNA technology has made an enormous impact on plant biotechnology, both in the development of novel crop traits and the functional analysis of new genes and their promoters. The efficient functional analysis of DNA fragments and the effective application of the resulting discoveries to crop trait improvement are increasingly dependent on innovative vector design and construction. The design and construction of vectors has an impact on the versatility of experimental systems and influences the public acceptability of genetically modified crops.

8.4.1. Vectors for High-Throughput Functional Analysis

Obtaining genomic DNA sequences for various monot and dicot plant species (an extensive compilation of plant genomic sequences can be found at Phytozome, a joint collaboration between the Department of Energy's Joint Genome Institute and the Center for Integrative Genomics (www.phytozome.net)) has presented new challenges in the production and analysis of recombinant DNA. Large numbers of promoters and genes encoded by these genomes have been discovered, but many remain uncharacterized, providing an incentive to design and construct vectors with the capacity for high-throughput functional analysis. Traditional ligase-mediated cloning is no longer a practical approach to facilitate the analysis of all the genes and promoters from these model organisms. Plant vectors compatible with Gateway® recombination cloning or Gibson Assembly end-linking cloning have been generated to aid these analyses. (An example comparing traditional restriction-ligation-based cloning with Gateway cloning is demonstrated in Figure 8.14.)

Recombination-compatible collections of plant open reading frames (ORFs; a sequence encoding a polypeptide) have also been generated. Trimmed ORFs lacking 5′ or 3′ UTRs (i.e., containing

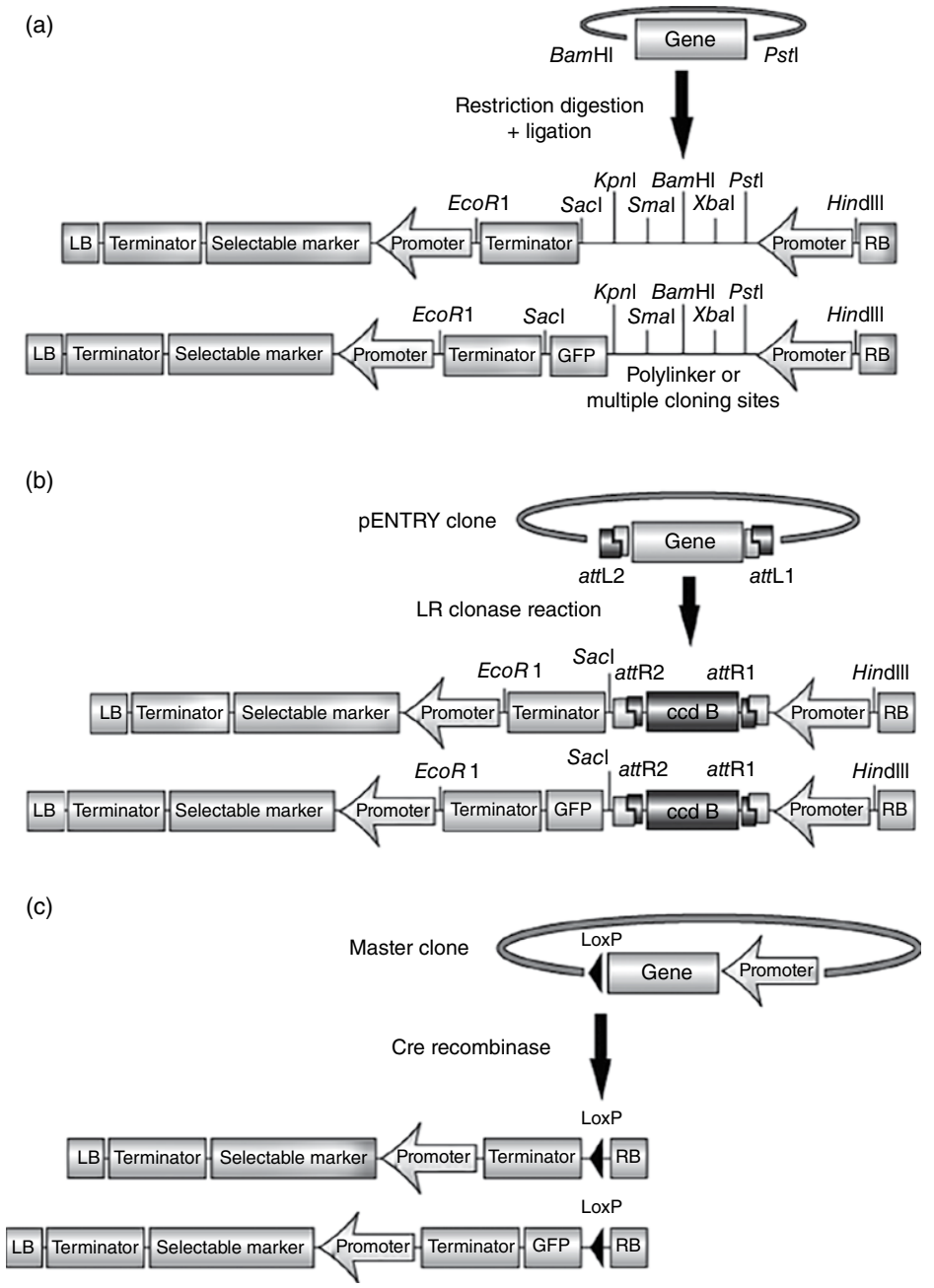


Figure 8.14. Plant gene expression vectors for conventional cloning using restriction digestion and ligation (a) and Gateway® recombination cloning (b). The first vectors shown in (a) and (b) are designed to allow a gene to be ectopically expressed in a plant cell. The second vectors shown for each category contain the green fluorescent protein (GFP) reporter gene. These vectors are designed to effect protein fusions with GFP to help identify the subcellular target of a protein under investigation. Ideally, three vectors for each type are frequently made, one for each reading frame, to ensure that a perfect fusion between the GOI and the marker gene is made. The insert DNA must be in an "open" ORF configuration (described in the text) so that no stop codon is present between the GOI and the marker gene.

protein-coding sequences only) can be shuttled rapidly and efficiently between vectors bearing compatible recombination sites. These so-called ORFeome collections have been generated so that the positions of the original translation initiation and termination codons remain intact ("closed" ORF configuration). However, since some applications to investigate gene function require the addition of C-terminal peptide fusions, ORFeome collections in which the stop codon is omitted ("open" ORF configuration) have also been generated. Often, the initial functional data on an ORF or gene are on the phenotype it induces when it is *ectopically* expressed (i.e., in tissues in which it is not normally expressed) under a constitutive and near-ubiquitous promoter. Gateway® vectors designed for this type of analysis have been generated using the strongly active 35S promoter from cauliflower mosaic virus (CaMV) for dicot species or the ubiquitin 1 promoter from maize (ZmUbi1) for monocot species. Some of these vectors have an additional design feature that provides stop codons adjacent to the 3′ recombination site in all three reading frames, to facilitate the expression of open as well as closed ORF configurations. Of course, not all ORFs can be misexpressed constitutively. Some cause lethal effects when expressed in this manner. In such cases, ORFs can be shuttled into vectors that are designed for conditional or inducible ectopic expression (Karimi et al. 2002; Curtis and Grossniklaus 2003; Joubes et al. 2004) or even to vector systems that allow induced expression in restricted cell types (Brand et al. 2006).

8.4.2. Vectors for Gene Down-Regulation Using RNA Interference (RNAi)

A very powerful tool that helps elucidate gene function is to reduce, or "knockdown" native gene expression in the organism using RNA interference (RNAi) (Waterhouse et al. 1998). Many different vector designs have been utilized to exploit RNAi for down-regulation of target genes, including virus-induced gene silencing (VIGS) (Thomas et al. 2001), self-complementary hairpin structures (hpRNA and shRNA—"short hairpin") (Waterhouse et al. 1998; Wesley et al. 2001), artificial microRNAs (amiRNAs) (Schwab et al. 2006), and synthetic trans-acting small interfering RNAs (tasiRNAs) (Montgomery et al. 2008; Carbonell et al. 2014). Some of these vector systems have been coupled with the Gateway cloning method. For instance, one of the earliest examples is the hpRNA system which contains double-stranded RNA produced by the transcription of an inverted repeat sequence of a gene. This transcript forms a hairpin-loop structure that triggers the RNAi pathway, leading to the degradation of homologous mRNAs (reviewed by Brodersen and Voinnet (2006)). The careful construction of specialized Gateway® destination vectors guarantees the rapid and efficient production of double-stranded RNAs (Fig. 8.15). In standard Gateway vectors, the *att* site modifications were designed to maintain DNA fragment orientation during the excision and integration process (Hartley et al. 2000). The arrangement of *att* sites in hpRNA constructs ensures the easy insertion of two identical gene segments in opposite orientations, downstream of a constitutively active promoter (Fig. 8.15). Constitutively expressed interfering RNA can be used to silence genes throughout a plant's development, or can be expressed conditionally to provide temporal control over the onset of gene silencing.

8.4.3. Expression Vectors

The thorough analysis of gene function frequently involves expressing a GOI, not only in plants but also in multiple systems. With traditional cloning methods, independently derived expression constructs must be made. Recombination cloning technology has revolutionized gene analysis by allowing genes to be expressed from the same recombination cassette in *E. coli, Saccharomyces cerevisiae*, or baculovirus expression systems (Liu et al. 1998) (Life Technologies), providing easier access to tools that broaden the scope for the functional analysis of genes.

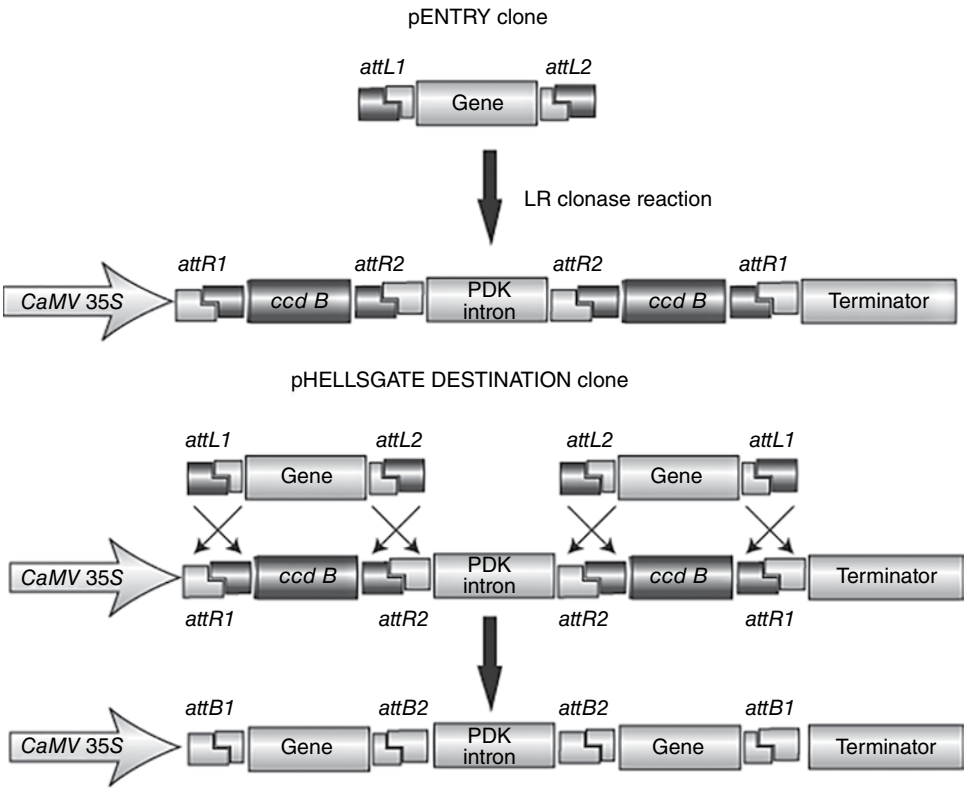


Figure 8.15. Gene silencing in plants can be achieved using inverted repeat transgene constructs that encode a hairpin RNA (hpRNA). Using Gateway® Cloning technology, the production of such inverted-repeat transgene constructs can be achieved efficiently, since DNA fragment orientation during the excision and integration process is maintained and the Gateway recombination cassettes are arranged in opposite orientations with respect to each other.

8.4.4. Vectors for Promoter Analysis

Tools that identify the spatial (where a gene is expressed) and temporal (when a gene is expressed) patterns of gene expression also provide important clues in functional genomics studies. Frequently, vectors are designed to allow the promoter or regulatory elements of a GOI to be fused upstream of a reporter-coding sequence (reporter genes are discussed in Chapter 10). Such constructs are used to determine the cell type(s), organ type(s), or developmental stage in which a gene is expressed. By assembling promoter ENTRY clones in recombination-compatible vectors, researchers are compiling a library of promoters and enhancers that are universally compatible with a wide variety of vectors. The modular assembly of DNA components has recently been extended through the introduction of additional novel recombination sites (Multisite Gateway®, Life Technologies) with unique specificities that allow multiple DNA fragments to be assembled in a single vector (Fig. 8.16). This facilitates the simultaneous incorporation of a promoter, ORF, and epitope tag into a single plant transcriptional unit (PTU) within a vector derived from collections of the modular component parts, or the simultaneous incorporation of multiple PTUs within the same vector backbone (referred to as a multigenic DNA construct).

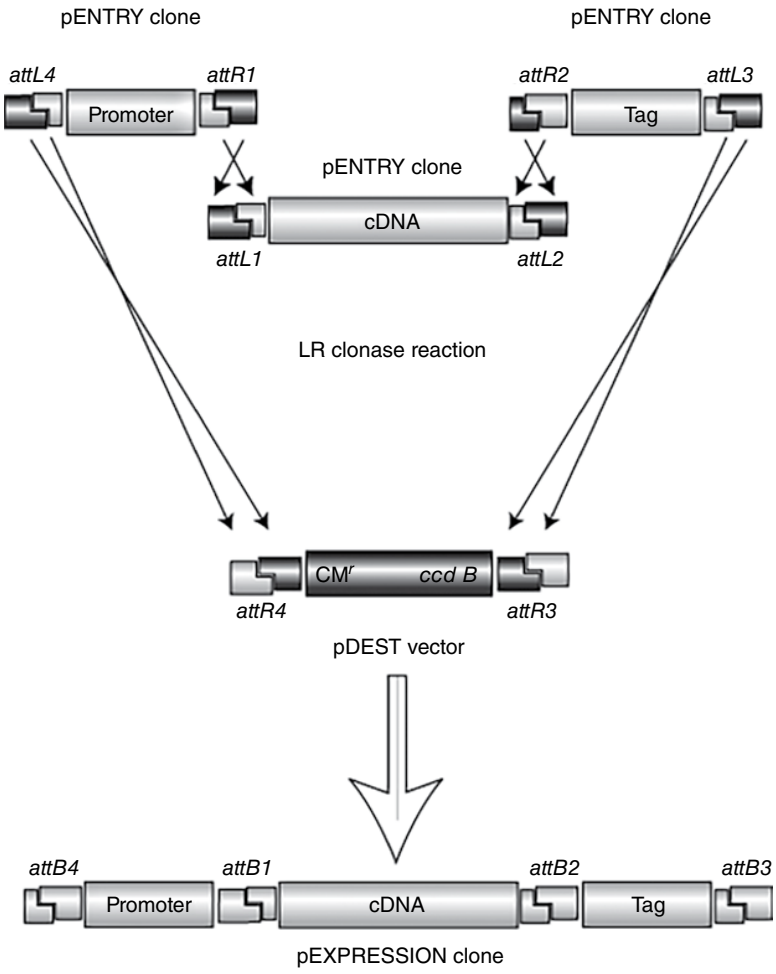


Figure 8.16. Multisite Gateway® allows several DNA fragments to be cloned within a single vector construct. More recent advances in the design of new *att* recombination sites have permitted the assembly of up to five DNA molecules within a single vector construct, but none have been designed as yet for plant transformation.

8.4.5. Vectors Derived from Plant Sequences

The revolutionary advances in recombinant DNA technology have resulted in significant scientific achievements, such as the replacement of animal-derived insulin with overexpression of human insulin in yeast and *E. coli* or the development of the recombinant hepatitis B vaccine. In plants, recombinant DNA technology provides opportunities to engineer new traits in crop plants that could not have been achieved through conventional breeding. However, consumer surveys have identified that public acceptance of genetically engineered organisms is often linked to concerns about the origin of the genetic material used to improve crop traits. These surveys have identified that the food crops least appealing to consumers are those containing foreign genetic material derived from organisms distantly related to plants. Ironically, wild-type plant cells already contain the genetic material of three genomes, the plant nuclear genome, and two bacterially derived genomes: the chloroplast and the mitochondrial genomes, from cyanobacteria and α-proteobacteria, respectively. Some researchers have responded to these concerns through the use of cisgenesis, which is the genetic modification of a plant with genetic elements (i.e., promoters, coding sequences, and terminators)

from within the sexually compatible group of species (Schouten et al. 2006). In contrast, transgenesis (i.e., production of transgenic plants) is defined as the modification of a plant with genes from a non-sexually compatible species (e.g., insect-resistant genes from *Bacillus thuriengensis*), synthetic genes (e.g., genes codon-optimized for plant expression), or genes arranged in combinations not found in nature (e.g., rice promoter combined with a microbial gene and a viral terminator). During the design stage of vector construction, measures can be taken to ensure that non-plant-derived sequences are kept to a minimum. The T-DNA of *Agrobacterium* is one source of foreign genetic material that could be eliminated using plant-derived "P-DNA" sequences (Rommens et al. 2004). These are functional analogs of *Agrobacterium-derived* T-DNAs, which have been shown to support the transfer of DNA from *Agrobacterium* to plant cells. Some examples of cisgenic plant production include apples with increased fungal resistance (Vanblaere et al. 2014) and barley with increased phytase activity in the grain (Holme et al. 2012). While cisgenic plants are limited in the origin of DNA that is introduced into the plant, they are under the same level of regulation as transgenic plants.

Since the transfer of DNA to plant cells is a relatively rare event, transformed cells are usually identified and regenerated with the aid of selectable markers, such as antibiotic resistance genes, traditionally derived from bacteria. Once these foreign selectable marker genes have served their purpose, they can be removed, since they play no further role in the expression of the transgenic trait. One method of removing such genes relies on the presence of an inducible recombination system in the plant vector, which allows excision of a marker gene positioned between recombination sites (Fig. 8.17). However, some marker genes are not removed, such as those conferring herbicide resistance, which can be used to select transformants in tissue culture and provide an economically important crop improvement trait in the field. In fact, about 90% of genetically

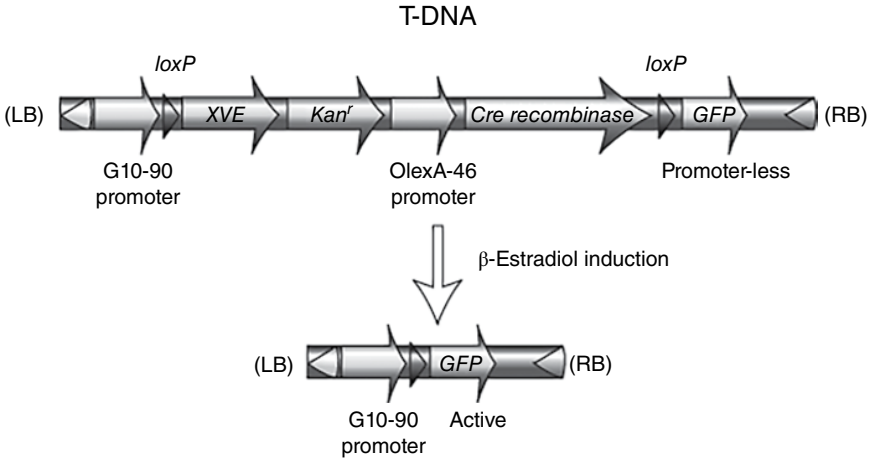


Figure 8.17. Excision of selectable marker gene following T-DNA insertion into the plant genome. XVE is a chimeric transcription factor. It contains three functional domains: a LexA DNA-binding domain (X), the VP16 activation domain (V), and the estrogen receptor-binding domain (E). The G10-90 promoter drives the constitutive and ubiquitous expression of XVE in transformed plant cells. The XVE protein is then bound as a monomer in the cytosol of the cell by a chaperone protein HSP90, and the target gene is transcriptionally inactive. Application of β-estradiol causes a conformational change in E, which leads to the release of HSP90 and dimerization of the receptor. On dimerization, the receptor is activated, allowing the protein to translocate to the nucleus of the cell where it binds O*LexA* binding sites of the promoter that is placed upstream of the Cre recombinase. The VP16 activation domain activates RNA polymerase II, leading to the transcription of the Cre recombinase gene. Cre recombinase allows recombination to occur between the *LoxP* sites removing all intervening genes, including the selectable marker gene.

modified crops are engineered for herbicide tolerance (Fernandez-Cornjeo 2014). Two of the most commonly used herbicide-resistant genes are derived from the bacteria *Streptomyces hygroscopicus* (Hoerlein 1994) and *Bacillus licheniformis* (Castle et al. 2004). Such bacterial herbicide-resistant genes could be replaced by plant-derived sequences. Several plant genes that produce agronomically useful levels of herbicide resistance have now been identified.

Plant genome sequence data and advances in plant molecular biology have provided the means by which to identify and isolate plant sequences that have the potential for use in crop improvement. Frequently, viral promoters, such as the CaMV 35S promoter, are used to express genes constitutively and near-ubiquitously in transgenic plants. These can be replaced by native plant promoters with similar expression profiles, such as actin or ubiquitin promoters. The use of such promoters to express transgenes both ubiquitously and constitutively may, however, cause unwanted secondary effects that might be avoided by designing and constructing vectors to deliver tissue-specific or conditional gene expression. For example, dwarfism is an agronomically important trait, which helps plants survive heavy rain and windy conditions. The dwarf plants of the so-called Green Revolution are short because they respond abnormally to the plant growth hormone gibberellin. Attempts to generate transgenic dwarf rice, by misexpressing the *Arabidopsis gibberellin-insensitive* (*GAI*) gene, resulted in short plants that unfortunately also produced low seed set (Fu et al. 2001; Tomsett et al. 2004). Subsequent experiments have shown that this problem could be resolved, at least in *Arabidopsis*, by constructing a vector that places the *GAI* transgene under the control of an inducible promoter (Ait-ali et al. 2003).

Although many endogenous (originating from within the organism) plant promoters that can rapidly respond to the application of inducers have been identified, these often also respond to environmental factors, such as water, salt stress, temperature, illumination, wounding, or infection by pathogens. Other nonendogenous inducible systems have been developed, but these rely on DNA sequences of foreign origin (for a review, see Curtis and Grossniklaus (2006)). Since endogenous promoters can be triggered inappropriately by environmental factors, and inducers may modify native gene expression (perhaps altering the physiology and development of the plant), an alternative approach that restricts transgene activity to specific tissue types to produce the desired trait would be more profitable. In the case of *GAI* expression, a construct with a promoter that is active in vegetative tissues only (and not reproductive tissues) may result in dwarf plants that do not have reduced seed set.

8.4.6. Vectors for Multigenic Traits

The construction of vectors for crop improvement can rely on the insertion of a single gene, as is the case with the production of *Bt* toxin to protect crops against insects, or on the insertion of several genes, as is the case with SmartStax®, a multitrait corn product that includes the consolidation of six separate insect-resistant traits and two separate herbicide-tolerant traits into one corn plant (Marra et al. 2010). Historically, multigenic traits were obtained either through sequential sexual crossing of transgenic plant lines that allows the accumulation of independent transgenes in a single plant (see Chapter 3), the successive transformation of transgenes into previously transformed plant material (Qi et al. 2004), or by the parallel introduction of different transgenes held on distinct T-DNAs into plants using co-transformation methods (see Chapter 11). The first two approaches mentioned are laborious, and the last one is technically challenging. Careful consideration of the design and construction of plant transformation vectors can resolve many of the technical difficulties, allowing multigenic traits to be expressed from a single T-DNA.

There are many alternative cloning approaches to "stacking" multiple genes into a single vector. These make use of the novel cloning methods discussed earlier (i.e., site-specific recombination systems, end-linking assembly methods and homing endonucleases) that allow the sequential

stacking of multiple expression cassettes into a single vector, thereby allowing the integration and expression of many transgenes from a single locus in the plant genome.

8.5. TARGETED TRANSGENE INSERTIONS

Once a recombinant T-DNA vector has been generated, with features designed to provide stable integration and gene expression, the DNA enters the plant cell and integrates randomly within the genome. The position of integration is uncontrolled and can often result in variable levels of transgene expression. A number of factors influence the level of transgene expression in plants, including the number of transgenes inserted into the genome, local *cis*-acting elements, and RNAi-mediated silencing. Nontranscribed, A/T-rich regions in eukaryotic genomes, known as *matrix attachment regions* (MARs), have been used to flank genes in T-DNA vectors (Butaye et al. 2004). These sequences have been reported to result in more reliable transgene expression shielding transgenes from RNA silencing (Mlynárová et al. 2003). However, targeting transgenes to predetermined chromosomal sites would perhaps provide greater control of gene expression and reduce potential positional effects. Until relatively recently, such approaches have been very inefficient in plants. Advances in the production of zinc-finger nucleases (ZFNs), transcriptional activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR)-Cas (CRISPR-associated) systems have now made it possible to increase the efficiency of targeted homologous recombination in plants through "designer" DNA target sequences within the plant genome (see Chapter 17). These systems have been demonstrated successfully for engineering precise deletions, insertions, or mutations within specific chromosomal regions (Cai et al. 2009; Petolino et al. 2010; Ainley et al. 2013). In addition, some of these synthetic molecules have been used as "designer" transcription factors for the regulation of gene expression within the plant (Petolino and Davies 2013, Liu et al. 2014). Customized nucleases and synthetic transcription factors provide a variety of precision tools to alter genomes and change the expression of endogenous genes and transgenes in future generations of genetically engineered plants (see Kim and Kim (2014) for a review and comparison of the ZFN, TALEN, and CRISPR/Cas systems).

Innovations in vector construction and plant transformation technology can influence the character of the resulting transgenic crop. One novel approach to reduce gene silencing and control transgene containment is the direct transformation of the plastid genome rather than the nuclear genome (Verma and Daniell 2007). This is because, unlike the nuclear genome, gene silencing does not occur in the plastid genome and, in most agronomically important plant species, plastids are maternally inherited, preventing pollen-mediated outcrossing. Vectors for chloroplast transformation are designed and constructed so that they contain left and right plastid-targeting regions (LTR and RTR), which are 1–2 kb in size and homologous to a chosen target site (Fig. 8.18).

The design and construction of plastid vectors that allow the simultaneous expression of several genes in an operon can be particularly useful in the engineering of agronomically important traits, as described earlier. Transgene integration has been achieved at 16 independent sites distributed across the plastid genome, ensuring that the positional effects, which are often associated with DNA integration events in the nuclear genome, are eliminated. Since there are 10–100 plastid genomes per plastid and approximately 10–100 plastids per cell, as many as 10,000 transgene copies can be generated in a single cell, resulting in highly abundant transgene transcription, producing as much as 40–50% of the total soluble protein in a cell. Plastid transformation technology does not yet extend to major crops, but has been demonstrated in numerous plant species including tobacco, soybean, carrot, and cotton through species-specific chloroplast vectors, and plant regeneration through somatic embryogenesis.

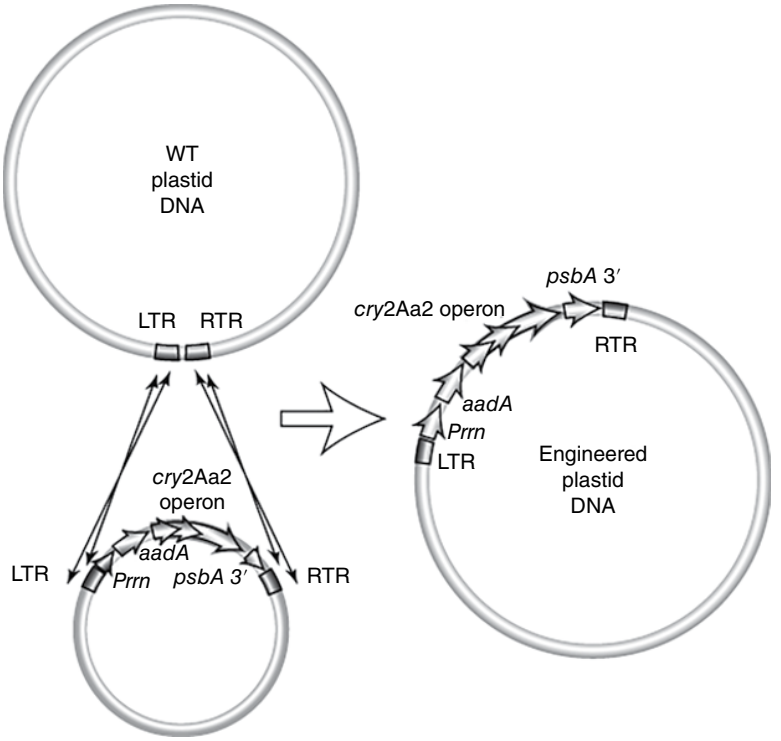


Figure 8.18. Site-specific integration is achieved by two homologous recombination events, one on either side of the DNA fragment to be integrated. During insertion, the targeted region of the vector replaces the targeted region of the plastid genome, and the vector backbone is lost. The inserted DNA fragment contains a selectable marker (here, the *aadA* gene encoding aminoglycoside 3′-adenylyltransferase, providing spectinomycin resistance) and can contain either a single gene flanked by independent 5′ and 3′ regulatory regions, including a promoter; a 5′ UTR and a 3′ UTR; or, as is the case here, multiple genes with a single promoter that regulates the expression of the operon with individual ribosome-binding sites (RBS), upstream of each ORF. In this example, the *Bacillus thuringiensis* (*Bt*) *cry2*Aa2 operon is inserted in the plastid genome generating insecticidal proteins in plant cells. The 16S ribosomal RNA (rRNA) promoter (*Prrn*) drives the expression of the *aadA* gene and the three genes of the *cry*2Aa2 operon. The terminator is the *psbA* 3′ region of a gene encoding the photosystem II reaction center component of the tobacco chloroplast genome.

8.6. PROSPECTS

Recombinant DNA technology, vector design, and construction form the foundations on which advances in modern plant biotechnology are built. The development of tools for the rapid amplification and manipulation of DNA sequences are essential if we are to keep pace the ever-increasing wealth of genetic information that results from the analysis of plant, animal, bacterial, and viral genomes. To exploit this information fully, functional studies must be conducted to determine the potential uses of such sequences, identifying the elements required to control gene expression and the genes required to ensure the high crop yields needed to sustain the planet's expanding population. Understanding the elements required for the efficient expression of genes in plants has already facilitated the development of new crop varieties. Novel genetic engineering approaches resulting from recombinant DNA technologies will provide the solutions to many of our future industrial, pharmaceutical, and biofuel requirements. This technology will continue to form the basis of our new "Green Revolution."

LIFE BOX 8.1. WAYNE PARROTT

Wayne Parrott, Professor, Department of Crop and Soil Sciences, University of Georgia.



Wayne Parrott. Courtesy of Wayne Parrott.

It was almost a given that I would have a career in agriculture—my parents were in agribusiness, and both sets of grandparents lived on a farm. Thus, upon graduating from high school, I started studying toward a degree in agronomy at the University of Kentucky.

It was not just the cultivation of plants that I found interesting—the plants themselves and their amazing diversity were just as fascinating—as was the underlying genetic basis behind all the diversity. Growing up in Central America, genetic diversity was all around me, as was access to "new and improved" varieties that breeders were releasing to farmers all the time. A key moment came while visiting my parents, who were living in Honduras at the time, and got invited to visit a banana-breeding station. Relating the experience to the Agronomy Club advisor back at Kentucky, he immediately offered me a job as a student worker in his wheat-breeding program. After that point, the genetic modification of crops—rather than their actual cultivation—became my chief interest.

The next key moment came when I saw my first plant growing in a test tube—it was so fascinating, there would never be any turning back. Coincidentally, the public press was filled with reports of the first gene transfer into a plant (the "sunbean") and all the expected potential to improve agriculture as that technology developed. I was irreversibly hooked.

Following graduation, I went to the University of Wisconsin for graduate school in plant breeding and genetics, where I got to indulge in studying all those aspects of plant genetics I found so fascinating. After graduation, I returned to the University of Kentucky as a postdoc. By that time, the foundation for plant tissue culture and genetic engineering technology had advanced to point it was a fruitful area of research. From there, I joined the faculty at the University of Georgia, where the technology and I have continued to grow up together.

Breakthroughs and major developments have not come continuously—they are interspersed with lots of experiments that don't work out or hypotheses that easily get disproved. Yet, when a technological hurdle is overcome, or when there is a new discovery—there is a rush and excitement that carries over to the next one. Looking back on my career, it is these moments that I most remember and collectively lead to a feeling of accomplishment.

But, research is also about people. The value of those who have served as my mentors along the way cannot be understated. I must mention Glenn Collins at the University of Kentucky, in whose lab I did an undergraduate research thesis, to whose lab I returned for a postdoc. Glenn has never stopped being my chief mentor. Then there is Richard Smith, my major professor from graduate school. I still see his work ethic and research approach in everything I do. Finally, I want to

single out Roger Boerma and Joe Bouton at the University of Georgia—two established faculty members who helped me out every step of the way.

Last but not least, I have had the good fortune to have had many postdocs, graduate students, and undergraduate students whose thought-provoking questions and enthusiasm have led my research forward. Seeing them move on and progress in their own careers has been as rewarding as the research itself.

LIFE BOX 8.2. DAVID MANN

David Mann, Research Scientist, Agronomic Traits Discovery Research, Dow AgroSciences.



David Mann. Courtesy of David Mann.

When I was young, I wasn't what you would call the "science kid." I didn't catalogue collections of leaves and insects. I didn't attend the NASA space camp or even build a volcano with baking soda. If asked what I wanted to be when I grew up, I gave the same answers my 5-year-old son now gives to me—a firefighter, a bridge builder, a detective, etc.—but not a scientist. After graduating from high school, I went off to college and majored in biology more by accident than with a detailed plan in mind. Needing to select a college major was mandatory for the scholarship I was applying for, so I checked the "Biology/Pre-Med" box on the application. Hesitant and uncertain in what I wanted to do with my life, I headed off to college with vague thoughts of being a medical doctor and the words of my dad lingering in my ear: "You can't steer a ship that doesn't move." But my thoughts of being a medical doctor promptly vanished when I realized during anatomy and physiology class that I didn't like the sight of blood. However, I resisted changing majors during my junior year when I heard a seminar speaker encourage floundering students to pick the hardest major they could find and stick with it. I stuck with it, eventually finding myself moving my tassel to the left side of my cap and learning that perseverance does pays off.

Following graduation, I had a couple of government lab internships at the US Army Institute of Chemical Defense (working in animal electrophysiology) and Oak Ridge National Laboratory (working with molecular biologists and engineers). I was now living in Knoxville, Tennessee and my wife was finishing up her undergraduate degree in social work. I was fortunate enough to be working with researchers at Oak Ridge National Laboratory who loved what they did. This included Guy Griffin and Tim McKnight. Tim was an energetic engineer and material scientist who had crossed over into biology and was developing and applying tools in nanotechnology to answer biological questions. His passion for his research and 1970s rock and roll music was contagious, and he was generous enough to ask questions and offer insight with my future career. The world of nanobiotechnology was exciting and new, and he introduced me to Gary Sayler and Mike Simpson, both professors at the University of Tennessee who were also

collaborating on the same research. I was graciously offered the opportunity to work on one of these projects with Tim and enrolled in a graduate program through the Department of Microbiology in Gary Sayler's lab. I worked with vertically aligned carbon nanofiber arrays, and using these tools, we were able to deliver DNA into single cells. The cells would then express the protein encoded within the DNA (we used fluorescent reporter genes), and then we could track the same cell over time and observe expression patterns. The technological tools were fun to work with, and during these years I grew in my depth of knowledge in molecular biology, biotechnology and microbiology, as well as my breadth of knowledge in engineering and material science. Additionally, the head of Microbiology persistently engrained in me the importance of hypothesis-driven research, especially when you are immersed in the field of biotechnology. These years were crucial to my development as a scientist, and I learned that even if you don't love something at first, sometimes the day-in and day-out practice over an extended period of time can cultivate a deeper passion for it (your practices help shape your passions).

With a Ph.D. in Microbiology and a focus on nanobiotechnology, I had many opportunities for postdoc positions in nanobiotechnology and synthetic biology labs. But I was more intrigued by the recent funding of the BioEnergy Science Center at Oak Ridge National Laboratory. The goal was to transform the field of bioenergy by demonstrating the feasibility of using renewable lignocellulosic feedstocks (poplar and switchgrass) as the source of sugar for production of biofuels. This was an exciting prospect, trying to reduce the recalcitrance of the plant cell wall (to break it down into usable sugars), while simultaneously maintaining the strength of the cell wall to preserve plant defense mechanisms and biomass production. The next 4 years in Neal Stewart's lab were an exciting time of applying the molecular biology experience I had previously gained, growing in my knowledge of plant transformation and plant physiology, and developing some useful tools for switchgrass (the bioenergy crop we were working with) along the way. I also learned to appreciate the importance of "cross-pollination" across many different fields of science to increase the scope and quality of a project. The project collaboration included molecular biologists, plant physiologists, agronomists, engineers, biochemists, plant breeders, computer scientists, microbiologists, ecologists, and more, all working together to come up with the best solutions to transform the bioenergy field. This inspired me in many ways and I had a larger view of ways that scientists in the plant research community were working with researchers in other areas to solve very big issues in the world at large as well as in the local communities we experience on a day-to-day basis.

Since moving on to the private sector (I now work for Dow AgroSciences), I have been able to apply these lessons from many amazing scientists who have taught me along the way and of which I am indebted to. I have grown in my passion (you might say I have caught the vision) for seeing how crucial science can be in solving the complex problems we are all now facing (i.e., how to feed the growing world around us). And while I am still young in my career, I love to see other "non-science kids" like me grow up to love the study of science as much as I now do.

REFERENCES

Ainley WM, Sastry-Dent L, Welter ME, Murray MG, Zeitler B, Amora R, Corbin DR, Miles RR, Arnold NL, Strange TL, Simpson MA, Cao Z, Carroll C, Pawelczak KS, Blue R, West K, Rowland LM, Perkins D, Samuel P, Dewes CM, Shen L, Sriram S, Evans SL, Rebar EJ, Zhang L, Gregory PD, Urnov FD, Webb SR, Petolino JF (2013): Trait stacking via targeted genome editing. *Plant Biotechnol J* **11**:1126–1134.

Ait-ali T, Rands C, Harberd N (2003): Flexible control of plant architecture and yield via switchable expression of *Arabidopsis* gai. *Plant Biotechnol J* **1**:337–343.

Brand L, Hörler M, Nüesch E, Vassalli S, Barrell P, Yang W, Jefferson RA, Grossniklaus U, Curtis MD (2006): A versatile and reliable two-component system for tissue-specific gene induction in *Arabidopsis*. *Plant Physiol* **141**:1194–1204.

Brodersen P, Voinnet O (2006): The diversity of RNA silencing pathways in plants. *Trends Genet* **22**:268–280.

Butaye KMJ, Goderis IJWM, Wouters PFJ, Pues JM-TG, Delauré SL, Broekaert WF, Depickers A, Cammue BPA, De Bolle MFC (2004): Stable high-level transgene expression in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. *The Plant J* **39**:440–449.

Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, Rock JM, Lee YL, Garrison R, Schulenberg L, Blue R, Worden A, Baker L, Faraji F, Zhang L, Holmes MC, Rebar EJ, Collingwood TN, Rubin-Wilson B, Gregory PD, Urnov FD, Petolino JF (2009): Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol Biol* **69**:699–709.

Carbonell A, Takeda A, Fahlgren N, Johnson SC, Cuperus JT, Carrington JC (2014): New generation of artifical microRNA and synthetic trans-acting small interfering RNA vectors for efficient gene silencing in *Arabidopsis*. *Plant Phys* **165**:15–29.

Castle LA, Siehl DL, Gorton R, Patten PA, Chen YH, Bertain S, Cho HJ, Duck N, Wong J, Liu D, Lassner MW (2004): Discovery and directed evolution of a glyphosate tolerance gene. *Science* **304**:1151–1154.

Clark AJ, Warren GJ (1979): Conjugal transmission of plasmids. *Annu Rev Genet* **13**:99–125.

Curtis MD, Grossniklaus U (2003): A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **133**:462–469.

Curtis MD, Grossniklaus U (2006): Conditional gene expression in plants. In *Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues*, Vol II, Teixera da Silva JA (ed). Global Science Books, London, pp 77–87.

Eguchi Y, Itoh T, Tomizawa J (1991): Antisense RNA. *Annu Rev Biochem* **60**:631–652.

Ellis T, Adie T, Baldwin GS (2011): DNA assembly for synthetic biology: from parts to pathways and beyond. *Integrative Biology* **3**:109–118.

Engler C, Kandzia R, Marillonnet S (2008): A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**:e3647.

Fernandez-Cornejo J (2014): United States Department of Agriculture, Economic Research Service, "Adoption of Genetically Engineered Crops in the U.S." Available at: http://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/recent-trends-in-ge-adoption.aspx (accessed on September 16, 2015).

Fu X, Sudhakar D, Peng J, Richards DE, Christou P, Harberd NP (2001): Expression of *Arabidopsis* GAI in transgenic rice represses multiple gibberellin responses. *Plant Cell* **13**:1791–1802.

Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison III CA, Smith HO (2008): Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**:1215–1220.

Gibson DG (2009): Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res* **37**:6984–6990.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison III CA, Smith HO (2009): Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**:343–345.

Hartley JL, Temple GF, Brasch MA (2000): DNA cloning using *in vitro* site-specific recombination. *Genome Res* **10**:1788–1795.

Hoerlein G (1994): Glufosinate (phosphinothricin), a natural amino acid with unexpected herbicidal properties. *Rev Environ Contam Toxicol* **138**:73–145.

Holme IB, Dionisio G, Brinch-Pedersen H, Wendt T, Madsen CK, Vincze E, Holm PB (2012): Cisgenic barley with improved phytase activity. *Plant Biotechnol J* **10**:237–247.

Joubes J, De Schutter K, Verkest A, Inze D, De Veylder L (2004): Conditional, recombinase-mediated expression of genes in plant cell cultures. *Plant J* **37**:889–896.

Karimi M, Inze D, Depicker A (2002): GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* **7**:193–195.

Kim H, Kim JS (2014): A guide to genome engineering with programmable nucleases. *Nat Rev Genet* **15**:321–324.

Kozak M (2005): Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**:13–37.

Li MZ, Elledge SJ (2007): Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat Methods* **4**:251–256.

Liu Q, Li MZ, Leibham D, Cortez D, Elledge SJ (1998): The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr Biol* **8**:1300–1309.

Liu W, Rudis MR, Peng Y, Mazarei M, Millwood RJ, Yang J-P, Yu W, Chesnut JD, Stewart CN Jr (2014): Synthetic TAL effectors for targeted enhancement of transgene expression in plants. *Plant Biotechnol J* **12**:436–446.

Marra MC, Piggott NE, Goodwin BK (2010): The anticipated value of SmartStax™ for US corn growers. *AgBioForum* **13**:1–12.

Mlynárová L, Hricová A, Loonen A, Nap J-P (2003): The presence of a chromatin boundary appears to shield a transgene in tobacco from RNA silencing. *Plant Cell* **15**:2203–2217.

Montgomery TA, Yoo SJ, Fahlgren N, Gilbert SD, Howell MD, Sullivan CM, Alexander A, Nguyen G, Allen E, Ahn JH, Carrington JC (2008): AGO1-miR173 complex initiates phased siRNA formation in plants. *Proc Natl Acad Sci USA* **105**:20055–20062.

Nordstrom K (1990): Control of plasmid replication—how do DNA iterons set the replication frequency? *Cell* **63**:1121–1124.

Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S (2005): Switches in bacteriophage lambda development. *Annu Rev Genet* **39**:409–429.

Patron NJ (2014): DNA assembly for plant biology: techniques and tools. *Curr Opin Plant Biol* **19**:14–19.

Petolino JF, Worden A, Curlee K, Connell J, Strange Moynahan TL, Larsen C, Russell S (2010): Zinc finger nuclease-mediated transgene deletion. *Plant Mol Biol* **73**:617–628.

Petolino JF, Davies JP (2013): Designed transcriptional regulators for trait development. *Plant Sci* **201–202**:128–136.

Qi B, Fraser T, Mugford S, Dobson G, Sayanova O, Butler J, Napier JA, Stobart AK, Lazarus CM (2004): Production of very long chain polyunsaturated omega-3 and omega-6 fatty acids in plants. *Nat Biotechnol* **22**:739–745.

Rommens CM, Humara JM, Ye J, Yan H, Richael C, Zhang L, Perry R, Swords K (2004): Crop improvement through modification of the plant's own genome. *Plant Physiol* **135**:421–431.

Schouten HJ, Krens FA, Jacobsen E (2006): Cisgenic plants are similar to traditionally bred plants: International regulations for genetically modified organisms should be altered to exempt cisgenesis. *EMBO Rep* **7**:750–753.

Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006: Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* **18**:1121–1133.

Sheng Y, Mancino V, Birren B (1995): Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Res* **23**:1990–1996.

Shetty RP, Endy D, Knight TF (2008): Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* **2**:5.

Thomas CL, Jones L, Baulcombe DC, Maule AJ (2001): Size constrains for targeting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using a potato virus X vector. *Plant J* **25**:417–425.

Tomsett B, Tregova A, Garoosi A, Caddick M (2004): Ethanol-inducible gene expression: First step towards a new green revolution? *Trends Plant Sci* **9**:159–161.

Vanblaere T, Flachowsky H, Gessler C, Broggini GAL (2014): Molecular characterization of cisgenic lines of apple "Gala" carrying the *Rvi6* scab resistance gene. *Plant Biotechnol J* **12**:2–9.

Verma D, Daniell H (2007): Chloroplast vector systems for biotechnology applications. *Plant Phys* **145**:1129–1143.

Waterhouse PM, Graham MW, Wang MB (1998): Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci USA* **95**:13959–13964.

Wesley SV, Helliwell C, Smith NA, Wang M-B, Rouse D, Liu Q, Gooding P, Singh S, Abbott D, Stoutjesdijk P, Robinson S, Gleave A, Green A, Waterhouse PM (2001): Constructs for efficient, effective and high throughput gene silencing in plants. *Plant J* **27**:581–590.

CHAPTER 9

Genes and Traits of Interest

KENNETH L. KORTH

Department of Plant Pathology, University of Arkansas, Fayetteville, Arkansas

9.0. CHAPTER SUMMARY AND OBJECTIVES

9.0.1. Summary

The whole purpose of biotechnology is to manipulate the genome of important plants, often by adding a few genes at a time. Traits can be manipulated by inserting DNA originating from any organism with that trait of interest into the target plant. Thus far in crop biotechnology, much work has been accomplished in conferring traits to plants such as the ability to survive herbicide treatment, insect resistance, disease resistance, and stress tolerance. However, there is growing interest in producing drugs and industrial proteins in plants as well as enhancing the nutrition of plant products.

9.0.2. Discussion Questions

1. What are the differences between "input" and "output" traits? Considering the environmental and biological factors that limit production in a farmer's field, what are some other input traits that might be good candidates for improvement using biotechnology?
2. Consider the possibility that you are employed by an agricultural biotechnology company and they ask you to find a bacterial gene for resistance to a specific herbicide. The herbicide has been manufactured by the company for many years. Using a strategy similar to that used to find glyphosate resistance, where might you start to look for a bacterium resistant to that herbicide?
3. What are the potential benefits of stacking multiple genes that confer resistance to one or more traits, such as herbicides and insects?
4. Golden Rice producing provitamin A has the potential to help many impoverished people who might benefit from eating it. Although application of this technology is supported by many people and organizations, there are also some who oppose the technology. Considering their possible motivations and potential biases, discuss some of the reasons that groups have come out in favor or in opposition to Golden Rice.
5. What are the potential benefits of producing pharmaceutical proteins in plants? What are some of the disadvantages or potential risks?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

6. Animal genes can be inserted into plants and expressed. Would you be opposed to eating foods expressing proteins encoded by animal genes? By human genes? What about if a transgenic plant produced a pharmaceutical? Discuss the reasons for your answers.

9.1. INTRODUCTION

As discussed in Chapter 6, the specific order of the nucleotide bases of DNA determines the function that a given sequence encodes. However, those four DNA bases are contained in a repetitious sugar-phosphate backbone that is essentially identical in DNA from any source. Because of this similarity of DNA structure in all organisms, there are no chemical limits on DNA from any organism being transferable to another, and this has allowed the development of transgenic plants carrying genes from many different sources, including microbes, insects, and animals, including humans. Essentially, sources for transgenes are as deep as our genomic knowledge in all of biology.

Many important traits in agriculture, such as crop yield, are often controlled by the action of multiple genes working together. However, other useful traits can be controlled by just a single gene. Because it has been easier to identify single-gene traits and produce transgenic plants with a limited number of introduced genes, most transgenic plants being grown today originated via the transfer of just one or a few foreign genes. In this chapter, the most common genes and traits that have been engineered in transgenic crops for improved agricultural production will be discussed, as will applications of transgenic plants that could benefit consumers by providing improved foods and products.

9.2. IDENTIFYING GENES OF INTEREST VIA GENOMICS AND OTHER OMICS TECHNOLOGIES

Advances in technologies for DNA sequencing and measuring mRNA accumulation have allowed detailed inquiry into the immense information contained in the genome of an organism. Genomics is a broadly defined term, but it generally refers to a strategy of using high-throughput, large-scale molecular techniques to analyze DNA sequence and function, or gene expression patterns.

Deciphering and interpreting the vast information of a genome sequence are the focus of great efforts, and it is hoped that this information will lead to development of new tools for crop improvement. In most crop species, this is a difficult task. For example, the soybean genome consists of around 1.1 billion base pairs (bp) of DNA, whereas the maize genome is considerably larger: approximately 2.4 billion bp. For comparison, the size of the human genome is slightly over 3 billion bp. These billions of base pairs of sequence are filled with many regions that are highly repetitive, and many others that do not seem to encode for any protein products. Identifying the important regions of plant DNA and those that contribute to useful traits for farmers can require a combination of traditional breeding techniques, high-tech molecular analyses, genetic studies, and newly developed computational strategies. The financial and intellectual commitments made toward completion of deciphering the human genome were instrumental in leading to development of new technologies for large-scale analysis of genes and proteins. Those technological developments continue today, and are being applied to analysis of every class of organism—including important crop plants.

Although all plant families and species have traits that make them unique, there are many genes that are conserved across species. In fact, there are many genes with conserved functions across plants and animals. By determining the function of a given gene in one species, it might allow us take a reasonable guess about the function of the corresponding, or "homologous," gene in another species. For this reason, some plants that are viewed as "model systems" get a lot of attention. For example, the species *Arabidopsis thaliana* is a small, fast-growing, member of the mustard family, and has a relatively small genome confined to just five chromosomes. For these reasons, it serves as

a good model for studies of plant development and response to environment. The *Arabidopsis* genome of approximately 120 million bp was the first of a plant to be fully sequenced (Arabidopsis Genome Initiative 2000). Having the complete genome of a plant, even one of no value as a crop such as *Arabidopsis*, has proven to be valuable for determining the function of individual genes. As new DNA sequencing technologies are developed and become cheaper, the genome information from crop plants increases and the similarities and differences among genes in different plant species are becoming clearer. It is hoped that by comparison of various crop genomes will lead to the identification of gene regions that are important for valuable traits.

From a technical perspective, new technologies have made it feasible to determine whole genome sequences of an organism (see Chapter 7). Although knowing the genomic sequence of a species is a valuable tool, it does not necessarily tell us about the function of genes or how they contribute to phenotype. It can be particularly difficult to associate specific genes with valuable traits, especially when the gene might have a minor, but important, effect on a trait. Therefore, genomic approaches to understanding gene functions or patterns of gene expression are being widely applied. Gene expression studies are typically aimed at indicating the presence of a particular mRNA transcript. For most genes, their ultimate function is dependent on the presence or form of the mRNA transcript whose nucleotide sequence information can be translated into amino acid sequence.

Expression of many genes is regulated at the level of mRNA accumulation and can be associated with their ultimate function in the plant. For example, genes thought to be involved in plant defense against pathogens will sometimes have greatly increased amounts of these gene-encoded mRNAs during infection. Scientists often take the approach to study this phenomenon by inoculating a plant with a pathogen, and then measuring mRNA transcript levels. If a given gene is "upregulated" at the level of mRNA accumulation, then this gene is a good candidate for being involved in defense responses. By measuring large numbers of transcripts under certain sets of environmental conditions, profiles of gene expression begin to emerge and gene sets involved in plant defenses (or other traits) can be identified.

One technique for measuring mRNA transcript accumulation of large numbers of genes is a DNA microarray (Alba et al. 2004). This technique takes advantage of the property of two nucleotide segments with complementary sequences to bind together, or hybridize. If one of the sequences is tagged with a label that can be measured, then the amount of binding can be quantified. In a DNA microarray, specific sequences are typically bound to a substrate such as a glass slide or chip on a small scale. Differing technologies allow for the binding of up to millions of individual sequences onto specific locations within areas as small as $1\,cm^2$. Generally, DNA sequences from a given species are produced or adhered on a microarray substrate, and then hybridized with labeled copies of mRNA (usually in the form of cDNA) from a specific tissue or after some treatment, such as pathogen inoculation. If a given mRNA is present at high levels, then a high degree of binding to its corresponding DNA sequence on the array will be detected. The level of binding of transcript sequences is usually compared with levels in some untreated control tissue. This general approach, known as comparative or differential gene expression, allows one to observe the transcript profiles of tens of thousands of genes in a single experiment.

A more widely used method for measuring mRNA profiles is a technique referred to as RNA Sequencing, or "RNA-Seq." This strategy takes advantage of advances in rapid and relatively inexpensive sequencing of large regions of DNA, sometimes called next-generation sequencing (NGS) (Egan et al. 2012). As opposed to the older microarray approach, this method does not require any prior knowledge about the genome sequence of the experimental organism or the development of microarray slides. This is a distinct advantage if one is studying the genome or gene expression in an agricultural plant that does not have a wealth of information about the genome. As with microarrays, comparative studies are often used in RNA-Seq experiments, and so any plant response to environmental treatments, or plant tissue, or genetic variety can be studied. Typically, mRNA is first isolated from the plant tissue(s) of interest and then converted to cDNA via reverse transcriptase. From this point, the cDNA is sequenced using the same techniques that are now widely applied to

sequencing of human and many other genomes. By comparing the mRNA sequences present in one experimental treatment versus another, it can be determined which genes are expressed or not at a given time or tissue, and thereby identify genes that might be important in conferring a valuable agricultural trait.

Ultimately, the protein products of most genes, or the metabolites that those proteins produce, are the things that lead to a particular plant trait of interest. It is therefore useful to analyze the end-products of gene expression. In fact, the accumulation of a given RNA transcript measured in most gene expression studies does not always correlate with the level or activity of the protein it encodes. This can be due to many factors, such as regulation of RNA stability, protein translation rates, or posttranslational regulation of protein stability or enzyme activity. As with genomic studies, the identification of an individual protein from among tens of thousands can be a technical challenge. Proteomic approaches use techniques to examine the large mixture of proteins present in a given tissue or after some treatment. This usually involves separating individual proteins on the basis of physical characteristics like size or charge. After the proteins are separated from one another, their amino acid sequence can be identified using techniques such as mass spectrometry. If the proteomic data are accompanied by a wealth of DNA sequence or gene expression data, it can be even more valuable, as the amino acid sequences can be correlated with specific gene sequences in that plant. Likewise, metabolomics is the large-scale analysis of chemical compounds that accumulate and contribute to the characters of a plant. These metabolites can be important for plant defense, physiology, nutrition, and food production; therefore, they are valuable contributors to a number of traits in crop plants that are of interest to farmers and consumers.

Through genomic, proteomic, and metabolomic approaches, scientists have attempted to take a large-scale, or "systems biology" view of the events occurring at the cellular level in an organism (see Chapter 7). The technologies developed and used in these methods generate massive amounts of data. Trying to make sense of these data is a considerable challenge in itself, and this has given rise to a discipline called "bioinformatics," which applies computational and mathematical methods to understand biological data (Rhee et al. 2006).

As the amount of genomic detail for crop plants continues to rapidly expand and be understood, it will provide more candidate genes as tools for biotechnological applications. Uses for this knowledge could come in the form of transgenes to be transferred between species, or as tools for plant breeders who utilize DNA marker-assisted selection in crop improvement. Importantly, as new tools for specific gene editing in plants are developed, many of these genes could be targets for directed modification to improve plant production or quality. This approach of specifically modifying genes already present in a crop species has been dubbed "cisgenic," as opposed to transgenic where genes are transferred from other species into a crop. The amount of information contained within a single plant species' genome is immense, and the potential that it holds for genetic improvement is therefore also large. Understanding and applying that potential is the challenge for scientists trying to identify genes that can contribute to traits of value to growers and consumers.

9.3. TRAITS FOR IMPROVED CROP PRODUCTION USING TRANSGENICS

In growing healthy plants that yield products of high quality and quantity, farmers must deal with ever-changing environmental conditions and pests. Transgenic approaches to helping farmers cope with these challenges are being broadly used today, while additional products are in the developmental pipeline. The most common uses of biotechnology in agriculture are engineered crops that are resistance to herbicides and pests. Plants with improved tolerance to abiotic stresses such as high temperatures, saline conditions, and drought are finding their way into production systems. Traits with more direct impact on consumers are improvements in grain and food quality that have been engineered via transgenic approaches. Therefore, it is clear that transgenic crops are addressing some of the oldest problems and issues faced in crop production.

9.3.1. Herbicide Resistance

The first transgenic application to be widely adopted in agriculture was resistance to herbicides. Weeds are generally regarded to be the most serious problem for farmers. Weeds compete with crops plants for water, light, and nutrients, thereby causing reduced crop yields. Chemical herbicides are widely used by farmers because they are cost-effective and efficient at killing weeds. Effective herbicides for agricultural production must be somewhat selective, meaning that they should kill the target weeds but not the crop plant. Using single-gene traits in transgenic plants can provide a very specific way to protect the crop plant from the effects of a given herbicide.

Herbicides generally work by targeting metabolic steps that are vital for plant survival. For example, glyphosate kills plants by inhibiting the production of certain amino acids that the plant requires for survival. Glyphosate is the active ingredient in the herbicide RoundUp™. Thus, crops such as soybean and corn that have been engineered to be resistant to glyphosate were given the name "RoundUp Ready."

Glyphosate works by binding to and inhibiting an enzyme called 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS), which is active in the shikimate pathway leading to the synthesis of chorismate-derived metabolites, including the aromatic amino acids (tyrosine, phenylalanine, and tryptophan) (Fig. 9.1).

To make plants resistant to glyphosate, a form of EPSPS that is functional in plants, but which is not affected by the herbicide was engineered into crops. In addition to being present in plants, the EPSPS protein can also be found in bacteria. So scientists at Monsanto, the makers of RoundUp, looked for and identified a form of EPSPS from a soil bacterium that was not sensitive to treatment

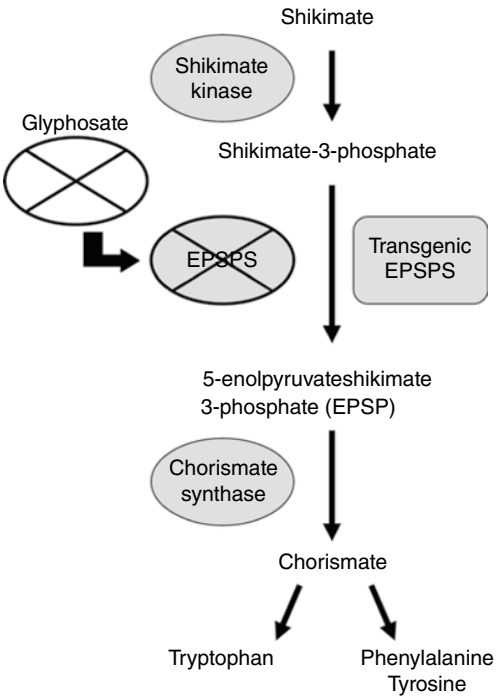


Figure 9.1. Resistance to glyphosate in RoundUp Ready™ plants is engineered by expressing a form of the 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase (EPSPS) enzyme that is resistant to the herbicide. In the absence of this transgenic enzyme, glyphosate inhibits the plant EPSPS and ultimately blocks the synthesis of chorismate, the branchpoint precursor to the essential aromatic amino acids: tryptophan, phenylalanine, and tyrosine. The transgenic EPSPS is unaffected by glyphosate, and can carry out the synthesis of EPSP leading to chorismate production.

with glyphosate. The initial steps in this process were relatively straightforward, as they simply plated soil bacteria on media containing glyphosate to identify strains that were resistant to the chemical. The EPSPS gene from the bacterium was then isolated and transferred into plants where its expression was regulated by putting it downstream of a strong promoter, the cauliflower mosaic virus 35S promoter, which drives gene expression throughout the plant (see Chapter 10) (Shah et al. 1986). Because *Agrobacterium*-mediated transformation methods are not very efficient in soybean, the particle bombardment method was used to make the initial transgenic event. This event was then used to transfer the glyphosate-resistant bacterial EPSPS gene to many other commercially grown soybean varieties using traditional breeding techniques.

The normal plant version of EPSPS is encoded by DNA in the nuclear genome. Following translation of the mRNA sequence to protein in the cytoplasm, EPSPS is transported into the chloroplast where the shikimate pathway is located. To ensure that the bacterial form of EPSPS would make its way into the chloroplast after the protein was synthesized, a short DNA sequence encoding a chloroplast transit peptide was fused to the 5′ end of the bacterial EPSPS open reading frame. This transit peptide sequence fused at the amino terminus of the bacterial EPSPS serves as an intracellular signal for proper protein localization. The transit peptide sequence originated from a gene encoding a protein normally found in the chloroplasts that carries out carbon fixation, ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Once the bacterial EPSPS gets into the chloroplast, it can function in place of the plant enzyme during the biosynthesis of aromatic amino acids when glyphosate is sprayed. Both the plant and engineered versions are resident in the engineered plant.

RoundUp Ready soybeans were one of the first transgenic crops to be approved by regulators and planted on a large scale (millions of hectares). Once they were commercialized, they gained rapid acceptance by farmers and are still the most popular transgenic plant in the world, and advanced versions of transgenic soybean exist now that are more effective at resisting the herbicide.

Glyphosate has several features that make it an attractive herbicide for growers. The compound is readily taken up and transported throughout the treated plant, features that make it especially effective as a herbicide. Because glyphosate is rapidly degraded by soil microorganisms, it does not persist long in the environment after application. This is a benefit both from an environmental standpoint and from a crop management perspective, because farmers can plant any crop in a sprayed field relatively soon after herbicide application. Because it is so effective at selectively killing weeds and not the herbicide-resistant crop plant, more farmers using glyphosate have adopted "no-till" or "low-till" methods, resulting in less soil erosion and lower fuel costs because they take fewer trips through a field (see Chapter 1). Furthermore, because animals do not make aromatic amino acids, they do not possess the shikimate pathway that is the target of glyphosate, and so the herbicide has low toxicity in animals.

In 1996, the first year they were commercially available, RoundUp Ready soybeans made up about 2% of the total soybeans grown in the United States. By 2000, that amount had risen to 54%, and in 2005 it was up to 87% (US National Agriculture Statistics Service; http://www.nass.usda.gov/). Now, glyphosate resistance has been engineered into a large number of crops that are grown all over the world, including Latin America and Asia. Predictably, adoption of glyphosate-resistant crops has resulted in a vast increase in the amount of this herbicide applied worldwide; however, there has been a decrease in the use of other herbicides, especially on soybean. This increase has also been encouraged by glyphosate coming off patent in 2001. Now glyphosate is sold as a generic herbicide by many companies as well in RoundUp formulations by Monsanto. The large amounts of glyphosate that are now applied to crops have led to concerns that glyphosate-resistant weed biotypes will be selected for and propagate in agricultural fields. Furthermore, farmers are required to pay a significant technology fee to Monsanto for the right to grow RoundUp Ready crops.

Recall that glyphosate resistance is conferred through the expression of an active target enzyme, EPSPS, which is not affected by the herbicide. An alternative strategy to express a protein that will inactivate an herbicide if it is sprayed onto plants. This is the approach used in resistance against the herbicide glufosinate, the active ingredient in the product Liberty™, generating a trait in crop plants

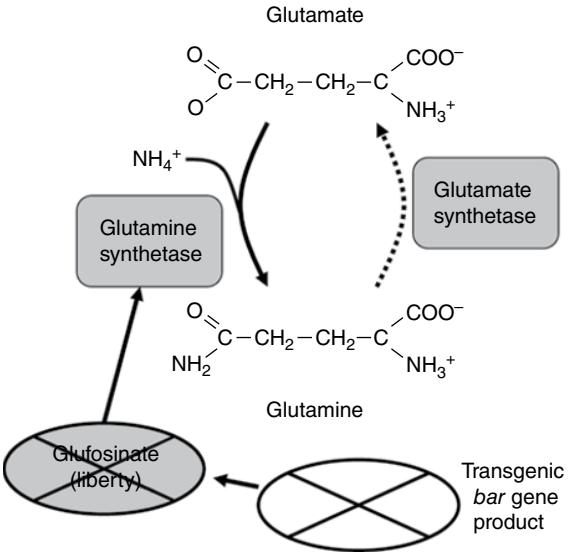


Figure 9.2. Resistance to glufosinate in LibertyLink™ plants is engineered by expressing an enzyme that directly targets and inactivates the herbicide. Glufosinate kills plants by inhibiting glutamine synthetase. This enzyme is responsible for production of the amino acid glutamine in a reaction that can sequester excess nitrogen by incorporating ammonia (NH_4^+). If this enzyme is inactivated by glufosinate, excess ammonia accumulates and the plant is killed. An enzyme encoded by the bacterial *bar* gene in transgenic plants inactivates glufosinate.

often called "Liberty Link." Glufosinate kills plants by inhibiting the plant enzyme glutamine synthetase (GS), which is responsible for synthesis of the amino acid glutamine. As part of the chemical reaction that produces glutamine, GS utilizes excess plant nitrogen in the form of ammonium which is incorporated into the amino acid. When GS is inhibited in glufosinate-treated plants, ammonium concentrations inside the plant rise to levels that are toxic (Fig. 9.2).

The glufosinate compound is naturally produced in some *Streptomyces* bacteria. In addition to having phytotoxic activity, glufosinate also serves as an antibiotic, because it is toxic to some other bacteria. Bacterial strains that are resistant to glufosinate produce an enzyme, encoded by the *bar* or *pat* gene, called phosphoinothricine acetyl transferase (PAT) (Thompson et al. 1987). The *bar* gene was isolated from a strain of *Streptomyces hygroscopicus*, which degrades glufosinate, and has been transferred into several crop plants. The Liberty Link trait is currently widely used in transgenic corn, canola, and cotton varieties.

The heavy use of herbicides in agriculture, with glyphosate perhaps being the most troublesome, has resulted in a very strong selective pressure for weeds that are resistant to the herbicide. In other words, weeds that were resistant to the herbicide survived and propagated, and now they cannot be controlled with glyphosate in farmers' fields. Obviously, this resistance eliminates the overall value for farmers to use glyphosate-resistant crops. There are currently dozens of problem weed species in the United States that have been characterized as glyphosate-resistant (see www.weedscience.org).

In response to the problem of glyphosate-resistant weeds, several companies are pursuing engineered resistance to alternative herbicides that have been in use since the 1960s. Similar to the strategy in making glufosinate-resistant Liberty Link crops, resistance to the herbicides dicamba and 2,4-dichlorophenoxyacetic acid (2,4-D) was engineered by expressing bacterial proteins that metabolize, and thereby inactivate, the herbicides. Both of these herbicides are plant growth regulators that are used primarily to control broadleaf weeds. Dicamba resistance is conferred by introduction of a gene that encodes a component of an enzyme complex called "dicamba monooxygenase." When the gene product is present, dicamba sprayed onto a plant is metabolized quickly so that it does not

impact plant growth. Likewise, a metabolizing enzyme is also used to confer 2,4-D resistance by modifying the herbicide if it enters the plant. It is hoped that by combining or "stacking" multiple herbicide resistance genes into the same crop plant, the effective use of herbicides will be prolonged many years by preventing strong selection for weeds resistant to a single herbicide. One prediction is that it will be much less likely for a weedy species to develop resistance to multiple herbicides with different plant targets, and because engineered crops will be resistant to several herbicides farmers can combine or alternate their weed-control treatments. One concern for growers is that both of these chemicals have previously been used in formulations that are somewhat volatile; and so if a farmer sprays his field, it could potentially impact neighboring fields because of the sensitivity of some plants to growth regulator herbicides. Agrichemical companies have developed forms of both 2,4-D and dicamba that are reportedly much less likely to have these "off-target" effects, but growers still need to exercise caution when applying these herbicides. Another concern is that many weeds have already evolved resistance to multiple herbicides (see www.weedscience.org) and that even dual use or alternating herbicides will not effectively address the weed resistance problem in the long run.

9.3.2. Insect Resistance

Insect damage to crops poses a problem for farmers all over the world. Despite great amounts of money and effort spent to control insect pests, staggering losses to insects are incurred before and after harvest. In an effort to control these pests, synthetic chemical insecticides are widely used where mechanized agriculture is practiced; but even so, insects continue as a significant hindrance to food production. In much of the world, insect damage proceeds unchecked by chemical pesticides, and growers and consumers suffer significant losses in both yield and quality.

A number of proteins with negative effects on insects have been tested as potential weapons to be used in engineering insect-resistant transgenic crops. Genes for several proteins have been expressed in transgenic plants and shown to inhibit insect growth or cause insect death. These include genes for a plethora of proteins, which include protease inhibitors, which interfere with insect digestion; lectins, which kill insects by binding to specific glycosylated proteins; and chitinases, enzymes that degrade chitin found in the cuticle of some insects. Although each of these genes has been shown to control insects when ingested, none have been as effective or widely adopted as the genes encoding endotoxins from the bacterium *Bacillus thuringiensis* (Bt). The natural insecticidal activity of Bt endotoxin proteins represents an attractive alternative to synthetic chemical pesticides, which often have non-selective toxic effects on beneficial insects, birds, fish, and mammals. The transgenic plant produces its own insecticidal protein that is delivered only to insects that dare eat the plant (Fig. 9.3).

Bacillus thuringiensis was first isolated and described over 100 years ago. A German microbiologist named Ernst Berliner formally named the species in 1915 when he found it killed flour moth caterpillars. His work followed and confirmed the discovery in 1902 of a bacterial disease affecting silkworms (*Bombyx mori*) in Japan. Obviously, infection by Bt is detrimental for silkworm production. However, it was later noted that Bt had toxic effects on caterpillar larvae of most Lepidoptera species (moths and butterflies), which gives the Bt species great potential as a tool for protecting crop plants. In later years, additional strains of Bt were identified that are toxic to Coleoptera (beetles), Diptera (flies and mosquitoes), and even nematodes, which, of course, are not insects. The specificity of insecticidal activity of Bt on a particular insect species is determined by the form(s) of the *Cry* gene(s) carried by the bacterium.

Only certain species of insects are controlled by particular endotoxins. The *cry* genes encoding the toxic proteins in Bt take their name from the *cry*stal inclusions formed inside the bacterium when it enters into its spore-forming stage. These crystals often contain more than one specific type of *cry* gene product. Before they become toxic, the *cry*-encoded Bt proteins exist as "pro-toxins" and must be activated inside the insect digestive tract. Once they are ingested by a susceptible insect, the crystals break down in the alkaline environment of the insect midgut, generally dissolving at pH 8.0 or greater.

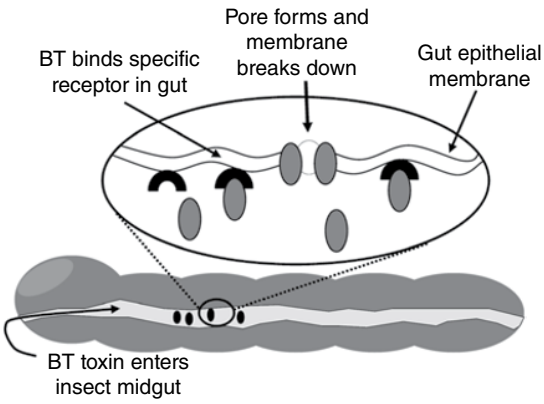


Figure 9.3. The Bt toxin binds to very specific receptors on the epithelial membrane of the insect gut. The toxin then forms channels in the membrane that leads to ion leakage and, ultimately, death of the insect. This mode of action explains the specificity of Bt (from the presence of the necessary receptors) and also shows why the toxin needs to be eaten by the insect to function.

At that point, the termini of the Bt pro-toxin proteins are cleaved by specific proteases inside the gut, yielding the toxic protein. The active protein will then bind to specific protein receptors on the insect microvillar membrane of the midgut (Fig. 9.3). In most cases, when Bt proteins are expressed in transgenic plants the entire coding region of the pro-toxin is not transferred to the plant. Rather, a shortened version of the gene will typically be expressed because levels of Bt protein accumulation are higher using this strategy (Barton et al. 1987; Fischoff et al. 1987).

After binding to a receptor the active Bt toxin will enter the insect cell membrane, where multiple copies of the protein will oligomerize and form pores. This results in ion leakage through the membrane, which causes membrane collapse from osmotic lysis. Once the membranes on the epithelia of the gut cells are disrupted, a massive infection or sepsis occurs and the insects effectively starve and die. In the case of a true *B. thuringiensis* infection, bacterial cells would form spores during the latter stages of infection and insect collapse, and thereby readying themselves for subsequent infections of other insects. In transgenic plants, susceptible insects usually stop feeding within a few hours after feeding on the plants, and die a short time later.

It is generally the presence or absence of specific forms of midgut receptors that determines whether a particular insect species is susceptible to a given Bt protein (Hofmann et al. 1988). For example, the most widely deployed *cry* genes in transgenic plants are members of the *cry*1A gene family, which are toxic to a broad range of Lepidoptera pests. However, this form of Bt has relatively little effect on Coleoptera species because the insects lack the specific receptors that recognize Cry1A proteins. Likewise some beetle species, such as the Colorado potato beetle (*Leptinotarsa decemlineata*), are targeted by the Cry3A Bt toxin whereas most lepidopterans are unaffected. Therefore, specific *cry* genes have been expressed in transgenic crops to tailor varieties to control specific pests and not affect nontarget species. For example, several variations of *Cry*1A genes have been transferred to corn to control European corn borer (*Ostrinia nubilalis*), a lepidopteran pest that feeds on the insides of corn stems; whereas *cry*3Bb1 expression has been used in corn varieties to control western corn rootworm (*Diabrotica virgifera*) larvae, a coleopteran species that feeds primarily on roots. By using this strategy, varieties resistant to a particular insect pest can be effective in growing regions where particular pests are problematic.

Because of the steps necessary to activate them and their target sites in the digestive tract, the Cry toxins are not effective as contact insecticides. Rather, insects are killed only when the toxins are ingested. This means that most nontarget and beneficial insects are not affected in fields of Bt crops. Furthermore, most insect and non-insect species lack the specific membrane receptors for Bt and often have digestive conditions that degrade the Bt toxin if it is consumed; therefore, Bt is essentially

non-toxic for most arthropods, animals, and birds. In fact, Bt sprays (the intact microbes) are considered to be so safe that certified organic food production in the United States allows for the direct application of Bt crystalline spores on plants immediately prior to harvest as a control for insects. Organic growers use Bt in this form as a valuable tool for insect control. One disadvantage of this approach in comparison to transgenic Bt production in plants is that Bt applied externally to plant surfaces does not penetrate the plant tissue and is not very stable, since it breaks down with time and exposure to ultraviolet light. Even so, because organic producers sometimes depend on application of Bt as a management tool, they are especially concerned about the possibility of the evolution of Bt-resistant insect populations caused by the widespread cultivation of engineered Bt crops.

As with herbicide-resistant crops, adoption of Bt transgenic crops has also been extensive. Damage by insects can be a severe problem in cotton, and this crop is heavily treated with synthetic chemical pesticides in many production schemes. In 2005, transgenic cotton represented almost 80% of the total of that crop grown in the United States, and it is widely grown in other parts of the world including China. Transgenic corn is now grown on approximately 90% of all the acreage in the United States. In the case of both cotton and corn, traits of herbicide resistance and insect resistance are often combined in the same plant lines as stacked traits.

As with the strong selective pressure that has led to problem weedy species in RoundUp Ready production systems, the widespread use of a single Bt gene means that insect species could evolve strong resistance to the protein. For this reason, and to utilize alternative modes of insecticidal action, other forms of engineered insect resistance have been commercialized and many others are being explored. For example, scientists again used bacteria as a source for an insecticidal gene product. The resulting vegetative insecticidal protein (VIP) also targets receptors in the insect gut, but these are independent from the Bt toxin receptors. As with stacking different types of herbicidal resistance in the same plant, the hope is that by stacking multiple forms of proteins, including dual Bts with different targets in insects, it will be much less likely for resistance to result in insect populations.

9.3.3. Pathogen Resistance

Pathogens such as viruses, fungi, and bacteria are severe and constant threats to crop production. Multiple transgenic approaches have been used to attempt plant disease control, although relatively few of these have made their way into the field of production.

The most effective way to control pathogens in a field setting is to use plants that are resistant to the problem pathogen. Resistance to a particular pathogen can often be conferred by a single plant gene (an *R* gene), the product of which is active in recognition of the presence or activity of a single virulence factor from the pathogen (encoded by an *Avr* gene). In plant-pathogen systems, this relationship is known as a gene-for-gene interaction (Fig. 9.4). Plant breeders have historically taken advantage of this system, although it can sometimes take many years to identify a plant line with the desired resistance and to breed that trait into useful cultivars. Another disadvantage to the breeding approach is that unwanted or undesirable genes may sometimes be linked to the *R* gene, and it can be difficult to separate them using traditional breeding methods. Finally, useful *R* genes are sometimes not easy to transfer because of barriers in crossing different species. Therefore, the ability to clone and transfer a single *R* gene from one plant variety or species to another represents an encouraging option to adapt and speed up the process.

A promising approach at engineering resistance is seen in the application of a specific resistance gene to ward off a bacterial disease in rice (Ronald 1997). Bacterial blight is a destructive disease of domesticated rice (*Oryza sativa*) in Africa and Asia, caused by the pathogen *Xanthomonas oryzae* pathovar *oryzae*. Scientists looking for alternative sources of resistance to bacterial blight identified a wild relative of rice, *Oryza longistaminata*, native to Mali, which is resistant to the pathogen but has very low grain quality and yield. Through careful genetic studies, an *R* gene called *Xa21* was isolated from the wild species. This gene has been introduced into domesticated rice using particle bombardment, and it confers strong resistance against strains of *X. oryzae* carrying the *Avr* gene recognized by *Xa21*.

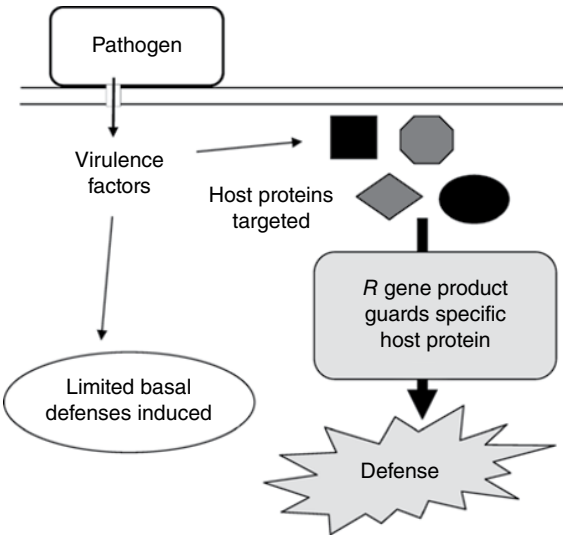


Figure 9.4. Resistance to specific strains of plant pathogens can be conferred by the protein product of a single resistance (*R*) gene. Most plant *R* genes function by recognizing the activity or presence of a specific virulence factor from the pathogen. In addition to the ability to induce basal defenses, these pathogen "effectors" are also active in attacking various host proteins. The protein products of *R* genes guard against pathogens via surveillance of specific targeted host proteins. When these *R*-gene-mediated defenses are triggered, the plant responds with a hypersensitive response and rapid activation of defense gene expression.

Through efforts of scientists scattered across the globe, the *Xa21* gene has been incorporated into several rice varieties of agricultural importance. The use of transgenic rice as a food crop is still controversial and its adoption has been slow compared to crops such as soybean, corn, and cotton. So although transgenic lines of blight-resistant rice are poised for application, they are currently not widely grown for food production. At least one-third of the world's population, including many developing countries, depends on rice as the major source of calories they consume. Therefore, development of disease-resistant rice could potentially make a major impact on alleviating hunger.

To date, viruses have been the most effective targets to control using transgenic plants. It has been known for decades that a previous inoculation with a virus can often protect a plant from subsequent infections by closely related viruses. This form of immunization of the plant has been known as cross-protection and has been employed with active viruses in limited cases. Crop plants can be intentionally inoculated with mild strains of a virus in the hope that this will protect the plant against future viral outbreaks. Much like vaccination with live viruses in humans, this strategy does have certain risks. In the case of inoculating with mild strains of a plant virus, there is a chance that the mild strain will present a drag on yield or that a virulent strain will emerge from the population and cause severe disease. With the advent of genetic engineering in plants, it became possible to express just a portion of plant viruses within the host. It turns out that this approach can likewise lead to resistance to closely related viruses.

Most plant viruses are relatively simple in terms of their genetic makeup, consisting of just a few genes carried by either an RNA or DNA genome encased in a protein coat. By expressing a portion of the viral genome constitutively in plants, a system of specific targeting of incoming, similar RNA sequences, can be activated in a potential host plant. This "RNA silencing" system is active in many organisms, including humans, and might have evolved partially as a surveillance and protection system against invading viruses.

A great success story using RNA-mediated virus resistance has developed in the production of papaya in Hawaii (Gonsalves 1998). Virtually, the entire production of this crop in Hawaii was threatened in the mid-1990s by the spread of the papaya ringspot virus (PRSV). The virus was so

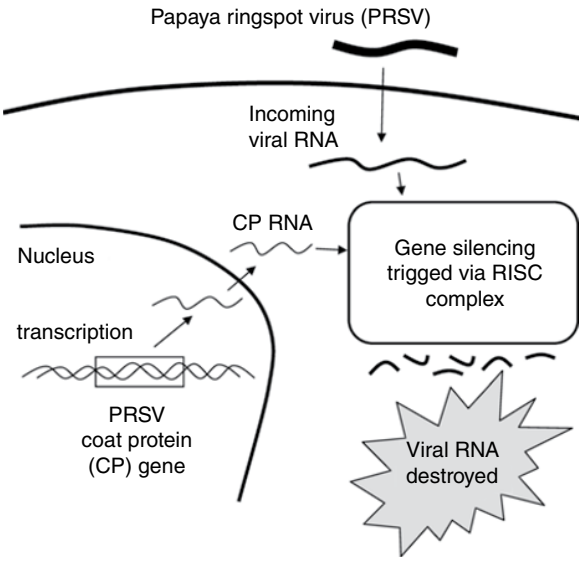


Figure 9.5. Transgenic resistance to papaya ringspot virus (PRSV) is possible because of the process of RNA-mediated gene silencing. To make virus-resistant plants, a portion of the coat protein (CP) gene of PRSV was transferred to and expressed in transgenic papaya plants. Following transcription, the RNA triggers targeted, sequence-specific degradation of similar RNA sequences, such as that found on incoming PRSV viral RNA. The initial degradation of RNA is carried out by an enzyme called *DICER*, and the process is mediated by an enzymatic structure called the *RNA-induced silencing complex* (RISC). Ultimately, this can lead to RNA cleavage, as well as blockage of transcription or translation of the target gene.

common, and the effects on yield were so severe that by the late 1990s many fields had been abandoned. By expressing the coat-protein gene of a mild strain of PRSV in papaya (Fig. 9.5), transgenic plants were made resistant to incoming pathogenic viruses (Fitch et al. 1992). Varieties of transgenic papaya were first introduced commercially in Hawaii in 1998; and so far, the transgenic lines have remained virus-resistant over the years. Just as in other transgenic crops, after the initial transgenic transformation in a single variety, the gene of interest was transferred to other desirable commercial varieties using standard breeding techniques. A similar approach has been used successfully to control cucumber mosaic virus (CMV) in transgenic squash production.

9.3.4. Traits for Improved Products and Food Quality

In the early years of commercialization of plant biotechnology, efforts and products focused on traits that help farmers in the growing of crop plants; these are called *input traits*. It is likely that many future applications of plant biotechnology will target *output traits*, centered on improved plant-based products that will find their way to consumers.

9.3.4.1. Nutritional Improvements. Humans depend on plants as food for survival. In addition to the calories that they provide, plants produce nutrients, vitamins, and essential amino acids that we require. Much more so than animals, plants have an incredible capacity for producing a variety of complex chemical compounds. Through methods in biotechnology, efforts are being made to take advantage of this capacity for chemical synthesis to improve or alter the nutritional values of plants.

One of the best known examples of nutritional improvement of a food crop has been the development of "Golden Rice," a transgenic plant that produces high levels of beta-carotene or provitamin A in the grain (Ye et al. 2000). Over one-third of the world's population depends on rice

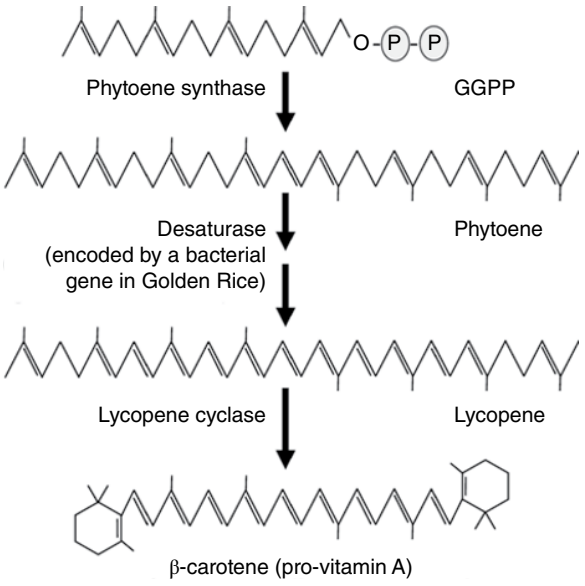


Figure 9.6. The production of β-carotene in Golden Rice was made possible by high-level, tissue-specific expression of the necessary enzymes in rice. Rice grains normally produce geranylgeranyl–diphosphate (GGPP). A gene-encoding phytoene synthase was transferred to rice from daffodil (for the original Golden Rice) or maize (in Golden Rice 2), and this led to production of phytoene in rice grains. A desaturase enzyme necessary to add double bonds to the structure was provided by transfer of a bacterial gene to rice (the two arrows at this step represent the multiple reactions that are necessary to add all double bonds). Finally, lycopene was converted in rice grains by an endogenous lycopene cyclase activity to the yellow-orange endproduct, β-carotene.

as a major component of their diet. Although rice can be a good source of calories, it is not high in protein or vitamins. Although dietary vitamin deficiencies are uncommon today in industrialized countries, they can still be a serious problem in developing countries in parts of southern Asia and sub-Saharan Africa, where rice is a staple and there is a lack of a diverse diet including meat, fruits, and vegetables. Vitamin A deficiency is especially serious, and the World Health Organization estimates that as many as 4 million children suffer with a severe deficiency. Humans depend on dietary sources of vitamin A, and deficiency of this vitamin is the leading preventable cause of blindness in children and makes sufferers significantly more likely to die from illnesses such as measles and malaria. Providing vitamin A supplements as capsules to children and new mothers is one approach to solving this problem, but to be effective supplements need to be administered several times per year, which can present logistical challenges in many areas. An alternative strategy is to provide provitamin A in the form of beta-carotene in rice.

Carotenoids are a subset of compounds within a large and variable class of plant metabolites called terpenoids or isoprenoids. This class of compounds is all based on a five-carbon building block, which can be assembled into multimers to form complex molecules. Many familiar plant scents and flavors, such as mint and pine resin, are based on terpenoids. The five-carbon precursor to terpenoids can be produced via two independent pathways, in either the cytoplasm or in plastids. Carotenoids are 40-carbon compounds produced from the precursor molecule via a biochemical pathway localized in plastids. The 40-carbon backbone of beta-carotene is phytoene, which is assembled by combination of two 20-carbon geranylgeranyl diphosphate (GGPP) molecules by the enzyme phytoene synthase (Fig. 9.6). Double bonds are then added to phytoene through a series of desaturation steps to produce lycopene, an antioxidant compound found in most plants and that contributes to the red color of tomatoes. Finally, lycopene can be converted to beta-carotene by the enzyme lycopene cyclase. Much of the understanding of how this pathway operates and could be

manipulated came from the laboratories of Dr. Ingo Potrykus in Switzerland and Peter Beyer in Germany. Researchers in these labs led the way in transforming rice with the necessary genes to produce carotenoids in rice grains.

Rice grains naturally produce GGPP, and so the addition of an active phytoene synthase gene expressed in rice grains under the control of a seed endosperm-specific promoter led to the production of phytoene in preliminary experiments. Transgenic plants were later produced via particle bombardment in which genes for phytoene synthase, phytoene desaturase, and a lycopene cyclase were co-transformed. These transgenic rice plants had grains with a bright yellow coloring, which was confirmed to come from the presence of beta-carotene and led to the name "Golden Rice" (Ye et al. 2000). It turned out that plants expressing just the phytoene synthase and the desaturase produced beta-carotene, indicating that rice grains already contained the metabolic activity to convert lycopene to beta-carotene. The gene for the desaturase originated from a bacterium, *Erwinia* spp., whereas the other genes came from daffodil. The bacterial desaturase enzyme actually performs metabolic steps normally carried out by two separate plant enzymes. Because the daffodil gene products are normally found in plastids, they already contained sequences for a plastid transit peptide to direct newly synthesized proteins to the proper cellular location. The bacterial gene encoding desaturase was modified by addition of a transit peptide to direct it to plastids following translation, in much the same way the bacterial EPSPS gene was modified for engineered RoundUp resistance.

Golden Rice produces carotene levels sufficient to impart a visible yellow color. One concern with these plants, however, has been that the accumulation levels of beta-carotene might not be sufficient to provide enough of the compound to be of optimal nutritional benefit. An improved version of transgenic rice referred to as Golden Rice 2, using a phytoene synthase gene from corn rather than daffodil, was subsequently produced that accumulated levels of carotenoids over 20 times higher than in the original Golden Rice (Paine et al. 2005). It is estimated that by eating modest amounts of Golden Rice 2, enough beta-carotene can be provided to overcome vitamin A deficiency.

The large-scale dissemination of Golden Rice has been controversial (see Life Box 9.2). Advocates maintain that this rice can provide provitamin A to millions of undernourished children who need it. Rice is already widely grown and consumed in the target regions, and so packaging the technology in this form takes advantage of an existing means to distribute and administer the nutrient. Opponents of the technology counter that development of this product is a tactic used by the biotechnology industry to drive acceptance of transgenic foods worldwide. Many opponents also contend that vitamin supplements and food fortification are superior methods for fighting the problem of vitamin A deficiency. Clearly, this rice has the potential to help malnourished children, but contentious issues will need to be resolved before it is accepted worldwide. At the very least, development of Golden Rice demonstrates that it is possible to alter the natural abilities of plants to synthesize complex chemicals, and to enhance their nutritional value.

9.3.4.2. Modified Plant Oils. The fatty acids produced by plants are the source of oils used in foods, and also have applications in cosmetics, detergents, and plastics. Oilseed rape (*Brassica napus*) has been used as a plant oil source for many years. Canola is the common name for the cultivated form of this plant, and has been bred through traditional means to contain low levels of harmful glucosinolates and erucic acid. By engineering canola with a thioesterase gene that originated in the California bay tree (*Umbellularia californica*), the oils that accumulate contain much higher levels of beneficial fatty acids. The "bay leaf" thioesterase enzyme expressed in canola causes premature chain termination of growing fatty acids, and results in accumulation of 12-carbon lauric acid and 14-carbon myristic acid. The overall level of lipids is not increased in these plants, as the increase in the short-chain molecules is matched by a decrease in the amount of long-chain fatty acids like the 18-carbon oleic- and linoleic acids. These short-chain fatty acids make the canola oil much more suitable as replacement for palm and coconut oils in products such as margarine, shortenings, and confectionaries.

Soybean oil is also used in a variety of food and industrial applications. By decreasing the levels of an enzyme called delta-12-saturase in transgenic soybeans, the amount of oleic acid can be increased. To decrease levels of enzyme expression, the normal soybean *fad2* gene encoding delta-12-saturase was repressed using a technique called "gene silencing," whereby a second copy of a portion of the gene is introduced into the plant where it forms a double-stranded (ds) RNA molecule. By overexpressing this dsRNA, a response in the plant is triggered to shut down expression of the endogenous gene. In this case, silencing the *fad2* gene results in higher levels of oleic acid and corresponding lower levels of two other 18-carbon fatty acids: linoleic and linolenic acids. The only differences in the structures of these three fatty acids are the number of double bonds in the chain. As a result, high oleic acid soybeans have low levels of saturated fats and trans-fats. This can alleviate the need for the hydrogenation process that is often used to make soybean oil suitable for foods like margarine, resulting in a healthier product. It also keeps the oil in a liquid form and makes it more heat-stable for cooking applications.

9.3.4.3. Modified Fruits and Vegetables. A promise of biotechnology has long been that consumers would benefit from improved nutritional and other qualities in their food. One example where this is the case is the production of engineered potatoes that produce lower amounts of a harmful compound. When potatoes are fried, such as for chips or French fries, the high temperatures cause accumulation of a compound called acrylamide, which is a known neurotoxin and a suspected carcinogen. These engineered Innate™ potatoes produce less of an amino acid, asparagine, which ultimately gets converted to acrylamide during the cooking process. As in the previous example for modifying oil biosynthesis, scientists used RNA-mediated gene silencing to decrease the expression levels of enzymes in the biosynthetic pathway leading to asparagine. Of much less dire consequences than acrylamide in your fries is the browning of an apple that occurs due to oxidation when you slice or bite into it. However, rapid browning is a trait that consumers do not desire, and its elimination might make possible the sale of sliced apples that do not require treatment with antioxidants. The transgenic Arctic Apple™ also utilizes gene silencing to confer a preferred trait. The Arctic Apple is engineered to have greatly reduced browning, achieved through silencing of the gene encoding polyphenol oxidase (PPO), which is directly responsible for production of brown oxidation products in the fruit. Although both Innate potatoes and Arctic Apple have been deregulated by the USDA so they can be grown for commercial use, it remains to be seen whether consumers will accept these products.

9.3.4.4. Pharmaceutical Products. The tremendous variety and potency of chemicals produced in plants has been long recognized, as many have powerful effects on human health and physiology (e.g., salicylic acid, cocaine, morphine, and taxol). In addition to being able to produce complex metabolites, plants can also produce high levels of specific proteins when a novel transgene is introduced.

Production of human and animal oral vaccines in plants has been proposed as an attractive approach, especially in areas of the world where infrastructure and costs might limit storage, transfer, and administration of traditional vaccines. By making an immunogenic protein in a food, vaccination could occur using a product that is easily grown and stored and that could be administered via consumption of the food source. For example, production of the surface antigen of the hepatitis B virus in transgenic potato has been demonstrated in clinical trials to lead to an immune response in humans consuming the potatoes. Production of proteins in transgenic bananas is also often cited as a potential source for these oral vaccines. There are several potential problems with this approach, such as the timing of administering the vaccine, dosage, and the ability of the protein to induce immunity upon oral administration. Nonetheless, this strategy might have application in some specific instances for humans or in vaccination of farm animals (Rybicki 2014).

Antibodies are large, complex proteins with the powerful ability to recognize and bind to specific molecular targets. Plants do not normally produce antibodies, but it has been repeatedly demonstrated that they can form functional antibodies when the encoding genes are expressed transgenically.

A high-profile example of this was the purification and first experimental use of antibodies produced in tobacco plants to treat humans infected with the often fatal Ebola virus in 2014. In this case, genes encoding mouse monoclonal antibodies that recognize the Ebola virus were isolated and their protein products produced following gene introduction into tobacco plants grown in a greenhouse. Although it was not clear if the product, called ZMapp, cured the patients tested, this episode clearly demonstrates the potential for using plants to manufacture high-value medicinal products.

To date, the vast majority of transgenic biopharmaceuticals are produced using *Escherichia coli*, yeast, or mammalian cell cultures. The strategy of producing pharmaceutical proteins in plants could have several advantages (Giddings et al. 2000). Transgenic plants offer the economies of scale to grow and harvest large amounts of biomass expressing the target product on relatively little land. Some applications for therapeutic proteins such as serum factors, hormones, or antibodies have traditionally relied on human or animal sources. By using plants, the risk of transferring unknown infectious agents from the donor source can be greatly reduced because plants typically do not carry animal pathogens. The idea of producing therapeutic proteins in crop plants is not accepted by everyone. Opponents worry that food products could be contaminated with tissue of plants intended for drug production. Another potential hurdle is the differences in glycosylation of proteins that occur in plants and animals. The sugar moieties added to proteins can vastly affect their function and immunogenicity, and some patterns of plant glycosylation can cause unwanted allergic reactions in humans. To be used in humans, these proteins would need to be produced so that they do not elicit an immune response in the patient.

9.3.4.5. Biofuels. With demands for energy increasing worldwide and supplies of fossil fuels being depleted, finding alternative and renewable energy sources has become an important goal for plant scientists. Both ethanol (ethyl alcohol) and biodiesel produced using plant materials can be adapted relatively easily to existing fuel storage, movement, and uses with existing infrastructure and machinery. Applications using transgenic plants have the potential to increase the efficiency of biofuel production on several fronts.

Ethanol offers several attractive features as an energy source; it is biodegradable, renewable, and burns cleaner than most fossil fuels. Ethanol is produced by yeast-driven fermentation of carbohydrates (sugars). In the United States, corn is currently the dominant source for fermentable sugars. In this case, the complex carbohydrates of starch in corn grains are first converted to simple sugars, which the yeast can then use to produce ethanol. One suggested approach to improve ethanol production is to transgenically engineer plants to produce higher levels of the enzymes responsible for the initial steps of starch breakdown (Himmel et al. 2007). The genes encoding enzymes like amylase, which degrades starch into simpler sugars, could possibly be expressed at high levels in corn grains or in other plants, resulting in higher percentages of readily fermentable sugars. The considerable inputs necessary for growing corn, in terms of nitrogen fertilizer, fuel, and pesticides, mean that it is likely not going to be an efficient long-term solution as a source for ethanol production. In Brazil, sugarcane is the plant source of choice for making ethanol, as the high levels of simple sugars make it superior for fermentation. In addition, sugarcane is a perennial crop that can be more easily grown with fewer inputs. The success of the Brazilian adoption of ethanol as a fuel source is widely touted as an example of how existing infrastructure and practices can be adapted for conversion to reliance on biofuels.

The use of plant material high in cellulose as a source for ethanol production is also being widely studied. The conversion of high-cellulose materials into fermentable sugars is an inefficient process, and so it is not currently viable as a method for biofuel production. However, plant materials such as corn stover (stalks and leaves), wood chips, or grasses contain energy that could potentially be converted to ethanol. Plant species such as the perennial switchgrass or fast-growing trees such as willow or poplar have advantages in that large amounts of biomass can be harvested multiple times from the same plants, and that they will grow efficiently with less need for watering and fertilizers. Although it is currently not efficient, improved methods for this "cellulytic conversion" of plant material to ethanol may hold some promise for sustainable fuel production from plants. Transgenic

approaches are being explored to produce cellulose that would be more easily used by yeast or other microbes for alcohol production, and decrease or alter lignin composition in plant cell walls. There have been single gene alterations for lignin biosynthesis that have led to plants with more easily degraded cell walls that produce increased biomass in the field (Baxter et al. 2014). In addition, the identification and engineering of microbes that can degrade lignin or more readily convert cellulose and sugars to ethanol is also being explored (Stephanopoulos 2007).

Diesel fuel made from plant material, biodiesel, can also represent an alternative to fossil fuels. Diesel currently accounts for approximately 20% of the fuel consumed for transportation in the United States; therefore, finding a renewable replacement could have a considerable impact on the need for oil throughout the world. Biodiesel is produced from oilseed crops such as soybean and canola, through a process called "transesterification." The properties of biodiesel are slightly different from petroleum-based diesel, but it can be used alone as a fuel or in a blend of the two types of fuel. Although there are currently no transgenic applications to improve biodiesel production in oilseed crops, the two major sources for biodiesel (soybean and canola) are most often grown as transgenic plants.

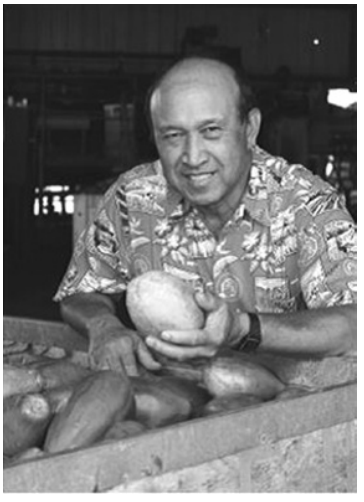
Because of the economic, environmental, and political concerns associated with fossil fuel consumption, the use of plants for biofuel production will almost certainly continue to increase and develop with new strategies.

9.4. CONCLUSION

Clearly, we are at the proverbial tip of the iceberg with regard to the numbers and types of genes identified that could be useful in plant biotechnology. Identification of candidate genes is limited by our knowledge of diverse genomes and by our understanding of the function of individual genes with the complex setting of a genome and whole organism. Simple solutions to problems that can be fixed with the insertion of one gene coding for one protein are myriad, but future advances will rely on the ability to engineer into plants with entire metabolic pathways such as was done to produce Golden Rice.

LIFE BOX 9.1. DENNIS GONSALVES

Dennis Gonsalves, Center Director, Pacific Basin Agricultural Research Center, USDA Agricultural Research Service; Recipient of the Alexander Von Humbolt Award (2002).



Dennis Gonsalves with transgenic papayas. Courtesy of Dennis Gonsalves.

I was born and raised on a sugar plantation in Kohala on the island of Hawaii. My dad was a first-generation Portuguese whose parents had immigrated from the Azores and from the Madeira islands. My mother was Hawaiian-Chinese with her dad emigrating from mainland China and her mom being a pure native Hawaiian. As a child and all the way through my undergraduate career, I never had ambitions to be a scientist nor even to go to graduate school. I had a key break in life when I was accepted to attend the excellent Kamehameha Schools, which had been started in the late 1800s by the Hawaiian Princess Pauahi Bishop to educate people of the Hawaiian race. I subsequently enrolled at the University of Hawaii with the intention of being an agricultural engineer so I could be back to work on the sugar plantation. However,

midway through my undergraduate tenure, the program for training engineers to work on the sugar plantations was dropped and I subsequently shifted to the field of horticulture. I was just an average student. I landed a job on the island of Kauai as a technician for Dr. Eduardo Trujillo, a plant pathologist at the University of Hawaii. That 1 year as a technician changed my life.

Dr. Trujillo told me to look at this "new" disease of papaya which he felt was caused by a virus. I knew next to nothing about viruses; but as soon as I started work to identify the disease, I knew that I wanted to be a research plant pathologist that would specialize in plant viruses. I had found my potential career niche. After working for some months as a technician, I wanted to pursue graduate work, but my grades were not good enough. I got a break when Dr. Trujillo persuaded the graduate school to accept me into a Master's program on probation. The other break or lesson also came from Dr. Trujillo who told me: "don't just be a test tube scientist, do things that will have practical applications." That philosophy would serve me well as I pursued my career, especially in biotechnology. I got my master's degree from the University of Hawaii in 1968 under Dr. Trujillo and Ph.D. from the University of California at Davis in 1971 under Dr. Robert Shepherd, who at that time had just shown for the first time that the cauliflower mosaic virus had a DNA genome. Little did I know that it would yield the sequences for the CAMV 35S promoter that is widely used in biotechnology. In 1972, I took a job at the University of Florida, subsequently moved to Cornell University in 1977, and in 2002 I returned to my Hawaiian roots to work for the Agricultural Research Service of USDA in Hilo, Hawaii

I gravitated from classical virology to molecular biology and biotechnology in the mid-1980s because of the prospects for developing virus-resistant transgenic crops. The pioneering work by Roger Beachy's group provided the proof of concept. My lab, in collaboration with others, have developed commercial virus-resistant squash and papaya. However, the papaya story has garnered the most interests for several reasons. A nutshell summary of the papaya work follows.

We developed, for Hawaii, transgenic papaya that resists papaya ringspot virus (PRSV), the most widespread and damaging virus of papaya worldwide. We started developing the transgenic papaya in the mid-1980s and had obtained a resistant transgenic papaya line by 1991. Coincidentally, PRSV invaded papaya plantations in Puna on Hawaii Island in 1992; and by 1995, the papaya industry was severely affected because 95% of Hawaii's papaya was being grown in Puna. Essentially, we had a potential technology to control the virus, but it had to be deregulated by APHIS and EPA, and pass consultation with FDA. We worked feverishly to test the papaya, develop data for deregulation, and get it commercialized. In 1998, we commercially released the SunUp and Rainbow papaya and essentially saved the industry from being devastated by PRSV. Nine years after commercialization, the transgenic papaya is widely grown in Hawaii and its resistance has held up well.

Aside from helping the Hawaiian papaya industry, our papaya work showed that "small" scientists can develop and commercialize a transgenic product. Basically, the work was done on a shoe-string budget and without funding from private companies. I and the team did the work because we were committed to help the papaya growers and to do it in a timely manner. If one analyzes the papaya story, one sees the ingredients for successful research and implementation because (a) work was done proactively by anticipating the potential damage that PRSV could do in Hawaii, (b) the research was focused so we could go from concept to practicality in a timely manner, (c) the research team had a strong commitment to good science and to achieving practical results in a timely manner, (d) the clientele was brought in and consulted early, and (e) we ventured out of our fields of expertise to get the job done. This last step involved collecting data needed for deregulation, assembling the package for submission to the regulatory agencies, working on the intellectual properties of the project, and making the clientele well aware of events as the project progressed. Today, the papaya case is often used as a model on how to get the job done in timely manner and make an impact, even though your group is small and your resources rather limited.

Plant virology is in an academic heyday, in part, because the technology of developing virus-resistant transgenic crops is now rather routine, and much is known about the mechanism that a governs resistance: post-transcriptional gene silencing. I expect to see continued incremental improvements on the development of effective virus-resistant transgenic crops. However, I am rather disappointed and surprised that so few transgenic virus-resistant crops (papaya, squash, and potato) have been commercialized. It is not due to lack of technology; numerous scientific reports have validated the effectiveness of virus-resistant transgenic plants with a number of plants and viruses. Yet, only transgenic squash and papaya are in commercial production today. We need to seriously ask why? Unless we effectively address this question, the huge promise that biotechnology has shown for virus-resistant crops will largely remain in the field of academia with little practical application. I suspect that the answers to this question do not lie in the technology arena, but more in the people's arena.

LIFE BOX 9.2. INGO POTRYKUS

Ingo Potrykus, Chairman, Humanitarian Golden Rice Board and Network; retired Professor in Plant Sciences, ETH Zurich.



Ingo Potrykus. Courtesy of Ingo Potrykus.

Rice-dependent poor societies are vitamin A-deficient because rice, their major source of calories, is totally devoid of provitamin A. Hundreds of millions of people in the developing world, therefore, do not reach the 50% level of the recommended nutrient intake (RNI) for vitamin A, required to live healthy lives. We developed "Golden Rice" to provide provitamin A with the routine diet. Even with rice lines containing modest concentration of provitamin A, a shift from ordinary rice to Golden Rice in the diet could save people from vitamin A malnutrition. Recent studies established that Golden Rice, if supported by governments, could save, at minimal costs, up to 40,000 lives per year in India alone.

How did I get involved in science and genetic engineering of plants? My connection to biology dates back to my childhood and is that of an old-fashioned naturalist. Ornithology is, after 65 years, still my major hobby. My interest in molecular biology began only when I was already around 40 years old. I got fascinated by the phenomenon of totipotency of somatic plant cells. Having an engineer's mind and being concerned about the problem of food-insecurity of poor people in developing countries, I could not resist of challenging that potential for contributing to food security—and this let me into a scientific career as pioneer in the area of plant tissue culture and genetic engineering. As research group leader at the Max-Planck-Institute for Plant Genetics, Heidelberg (1974–1976), the Friedrich Miescher Institute Basel (1976–1986), and full professor in plant sciences at the Swiss Federal Institute of Technology (ETH) (1986–1999). I had exceptional good conditions and great teams to follow the basic concept of developing genetic engineering technology for crop plants such as cereals and cassava. The task was to rescue harvests, to improve the nutritional content, and to improve exploitation of natural resources. As long as active in academia this was all "proof-of-concept" work. Only with my retirement in 1999 and the need for Golden Rice to be brought to the poor did I realize that the decisive follow-up steps of product development

and deregulation are routinely ignored by academia.

The science leading to Golden Rice. By the end of the 1980s, we had transformation protocols ready for rice and had already worked on insect-, pest-, and disease-resistance. The Rockefeller Foundation alerted me of the problem of micronutrient malnutrition. In 1991, I appointed a PhD student to work toward provitamin A-biosynthesis in rice endosperm and Dr. Peter Beyer, an expert in terpenoid biosynthesis from the University of Freiburg, Germany, joined as co-supervisor. The project was, for numerous good reasons, considered totally unfeasible and was, therefore, difficult to finance. The breakthrough came 8 years later, with the concluding experiment of a Chinese postdoc. When the harvest from a co-transformation experiment involving five genes was polished, the offspring from a transgenic line harboring all genes segregated for white and yellow endosperm. This was in February 1999 and 2 months before my retirement.

There was great recognition but no support for completion from the public domain. We presented our success to the public at my ETH-Farewell Symposium on March 31, 1999, and it was finally published in *Science* (see Ye et al. (2000)). Nature refused our earlier submitted manuscript for publication because "of lack of interest." The scientific community, however, the media and the public were quite interested, even became excited about this vitamin A rice. TIME Magazine devoted a cover story on Golden Rice July 31, 2000, and there were hundreds of articles and airings in the media. The readers of *Nature Biotechnology* voted Peter Beyer and me as the "most influential personalities in agronomic, industrial, and environmental biotechnology for the decade 1995 to 2005" (Jayaraman et al., 2006), and there were numerous recognitions for the work from the scientific community and the public. However, nowhere in the public domain could we find support for the long and tedious process of product development for our humanitarian Golden Rice project.

The private sector helped us to continue with the humanitarian project. Only thanks to the establishment of a public–private partnership with Zeneca/Syngenta could we proceed. The basis was an agreement, in which we transferred the rights for commercial exploitation in return for their support for the humanitarian project—making Golden Rice freely available to the poor in developing countries. This public–private partnership also helped to solve the next big problem: getting permission to use all intellectual property rights involved in the technology. We had been using intellectual property of 70 patents belonging to 32 patent holders! Fortunately, 58 patents were not valid in our target countries; and of the remaining 12, 6 of these belonged to our partner company and for the rest, it was not a big problem to get free licenses. Product development, deregulation, and delivery of a GMO-product turned out to be a gigantic task, especially for two naive university professors. We needed advice from the private sector and received help from Dr. Adrian Dubock who worked for Syngenta. We were short in different areas of expertise for strategic decisions and established a Humanitarian Golden Rice Board. We needed GMO-competent partner institutions in our target countries and established a Golden Rice Network. And we needed managerial capacity and found a project manager and a network coordinator. For more details, please visit our www.goldenrice.org homepage.

Lost years because of over-regulation. If it were not a GMO, Golden Rice would have been in the hands of the farmers since 2003. We have lost more than 10 years in the preparatory adoption to regulatory requirements, which do not make any sense, scientifically. An example for how irrational regulatory authorities operate could be found in our experience obtaining permission for small-scale field testing of Golden Rice. No ecologist around the world has been able to propose any serious risk for any environment from a rice plant containing a few micrograms of carotenoids in the endosperm (in addition to large amounts in the leaves), and this trait does not provide for any selective advantage in any environment. The first instance for permitting field testing in India was granted in 2015! The Philippines were, fortunately, a bit faster (2009). Because of the requirements of the established "extreme precautionary regulation" in costs and data, we have to base all Golden Rice breeding work and variety development on one single selected transgenic event. The decision for

the "lead event" we took in 2010 although we had to decide without any agronomic data—an impossible decision! From then on, our partner institutions have been introgressing the transgenes from this event into 30 carefully chosen popular southeastern Asian rice varieties. Hypothetically, deregulation can then be based on the single event, not on 30 different varieties. Considering the 40,000 lives, Golden Rice could save in India per year, regulation is, through the delay it is causing, responsible for the death and misery of hundreds of thousands of poor people. And did regulation prevent any harm? Judging from all regulatory review and from all data from all "biosafety research," my answer is: "most probably not."

Where is plant biotechnology going in the future? The answer depends entirely upon what our society does with GMO regulation. If this unjustified and excessive procedure is maintained, plant genetic engineering will have no future, and hundreds of millions of lives will be lost, which could be saved by applying this technology to food security problems. Plant molecular biology is, so far, an extremely successful scientific discipline, but much of its motivation and funding came from its potential application, not only in the private sector, but much so in the public sector, for example as contribution to the solution of humanitarian problems. With this potential being cut off, financial support for basic research will probably dry out.

I propose here my recommendations needed for humanity to maximally benefit from plant biotechnology:

- De-demonize GMO's and inform the public that these are perfectly normal plants. There is scientific consensus that genetic engineering does not carry any inherent risks. And there is the simple fact that there is not a single documented case of harm from the use of the technology. And there is no doubt that there is not a single crop plant that has not been extensively "genetically modified" by traditional interventions.
- Reform GMO regulation such that it evaluates traits, not GM technology, and takes decisions on balancing benefits versus risks. Because of the time and financial requirements of present regulation, no public institution can afford to take a single transgenic event to the marketplace.
- Establish public funding schemes for product development and deregulation. Humanitarian problems are problems of the public sector and should not be expected to be solved by the private sector.
- Encourage public–private partnerships for the solution of humanitarian problems. The private sector has the experience necessary for solutions of practical problems.
- Establish a reward system for those in academia who sacrifice their academic career by contributing to solutions of humanitarian problems. Academia receives much of its funding because the public believes that it is helping to solve humanitarian problems.
- Change the "highest priority to biosafety" paradigm. It leads to millions of deaths, and there are other topics deserving higher priority such as food security and poverty alleviation.
- Prosecute those organizations who use their political and financial power to block green biotechnology in an international court. They are responsible for crimes against humanity.

REFERENCES

Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JK, Martin G, Tanksley SD, Bouzayen M, Jahn MM, Giovannoni J (2004): ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. *Plant J* **39**:697–714.

Arabidopsis Genome Initiative (2000): Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**:796–815.

Barton KA, Whiteley HR, Yang N-S (1987): *Bacillus thuringiensis* delta-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to lepidopteran insects. *Plant Physiol* **85**:1103–1109.

Baxter HL, Mazarei M, Labbe N, Kline LM, Cheng Q, Windham MT, Mann DGJ, Fu C, Ziebell A, Sykes RW, Rodriguez Jr M, Davis MF, Mielenz JR, Dixon RA, Wang Z-Y, Stewart Jr CN (2014): Two-year field analysis of decreased recalcitrance transgenic switchgrass. *Plant Biotechnol J* **12**:914–924.

Egan AN, Schlueter J, Spooner DM (2012): Applications of next-generation sequencing in plant biology. *Am J Bot* **99**:175–185.

Fischoff DA, Bowdish KS, Perlak FJ, Marrone PG, McCormick SM, Niedermeyer JG, Dean DA, Kusano-Kretzmer K, Mayer EJ, Rochester DE, Rogers SG, Fraley RT (1987): Insect tolerant transgenic tomato plants. *BioTechnol* **5**:807–813.

Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1992): Virus resistant papaya derived from tissues bombarded with the coat protein gene of papaya ringspot virus. *BioTechnol* **10**:1466–1472.

Giddings G, Allison G, Brooks D, Carter A (2000): Transgenic plants as factories for biopharmaceuticals. *Nat Biotechnol* **18**:1151–1155.

Gonsalves D (1998): Control of papaya ringspot virus in papaya: a case study. *Ann Rev Phytopathol* **36**:415–437.

Himmel ME, Ding SH, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007): Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**:804–807.

Hofmann C, Vanderbruggen H, Höfte H, Van Rie J, Jansens S, Van Mellaert H (1988): Specificity of *Bacillus thuringiensis* delta-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midguts. *Proc Natl Acad Sci USA* **85**:7844–7848.

Jayaraman KS, Louët S, Powell K, Ransom J, Sheridan C, Vastag B, Waltz E (2006): Who's who in biotech. *Nat Biotechnol* **24**:291–300.

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, Drake R (2005): Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol* **4**:482–487.

Rhee SY, Dickerson J, Xu D (2006): Bioinformatics and its applications in plant biology. *Ann Rev Plant Biol* **57**:335–360.

Ronald PC (1997): Making rice disease-resistant. *Sci Am* **277**:100–105.

Rybicki EP (2014): Plant-based vaccines against viruses. *Virol J* **11**:205.

Shah DM, Horsch RB, Klee HJ, Kishore GM, Winter JA, Tumer NE, Hironaka CM, Sanders PR, Gasser CS, Aykent S, Siegel NR, Rogers SG, Fraley RT (1986): Engineering herbicide tolerance in transgenic plants. *Science* **233**:478–481.

Stephanopoulos G (2007): Challenges in engineering microbes for biofuels production. *Science* **315**:801–804.

Thompson CJ, Movva NR, Tizard R, Crameri R, Davies JE, Lauwereys M, Botterman J (1987): Characterization of the herbicide-resistance gene bar from *Streptomyces hygroscopicus*. *EMBO J* **6**:2519–2523.

Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000): Engineering provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science*, **287**:303–305.

CHAPTER 10

Promoters and Marker Genes

WUSHENG LIU AND C. NEAL STEWART, Jr.

Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

BRIAN MIKI

Agriculture and Agri-Food Canada, Ottawa, Canada

10.0. CHAPTER SUMMARY AND OBJECTIVES

10.0.1. Summary

This chapter is about two essential segments of DNA found in transgenic plant vector constructs. (a) A *promoter* must be fused upstream of the gene of interest (GOI) and other genes in the vector to control their expression. The choice of promoter is crucial in that it specifies when and where a transgene is expressed in the plant. (b) *Marker genes* are needed to select transgenic plants and/or monitor gene expression. There are two types of marker genes used in plant transformation. *Selectable markers* typically confer antibiotic or herbicide resistance so that transgenic cells, tissues, and plants can be selected for survival when grown with an accompanying chemical that otherwise is toxic to the plant. *Reporter genes* often will cause a visible color change in the transgenic plants so that researchers can see when and where transgenes are expressed in plant tissues.

10.0.2. Discussion Questions

1. Why is the choice of promoter important?
2. Why do promoters work conditionally, spatiotemporally, or inducibly?
3. What are synthetic promoters, and how are they constructed?
4. What are some differences between selectable marker genes and reporter genes?
5. Discuss the relative merits of GUS and GFP or other FPs as reporters. Does the profile of experimentation using these reporter genes overlap directly or partially?
6. What are the advantages, if any, for the use of the *manA* gene over the *nptII* gene as a selectable marker for food and feed crops, and would the use of the *manA* gene overcome public concern over the use of the *nptII* gene? Conversely, what are the disadvantages?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

7. Considering the large number of selectable marker gene systems that have been developed, why are so few adopted for basic research and commercialization?
8. What experimental factors should be considered for a functional genomics study of unknown genes if the vector employs a new selectable marker gene system in the base vector?

10.1. INTRODUCTION

The genetic transformation of plant cells has been known to occur in nature for a long period of time. However, the technologies for genetic engineering of plants in the laboratory were developed in the early 1980s. Then, the first stably transformed plants—in this case, tobacco plants—were produced using *Agrobacterium tumefaciens*, which can deliver a segment of its DNA (i.e., transfer DNA or T-DNA) into the plant genome. The "binary" plasmid is the plasmid that contains the engineered T-DNA for trait delivery. The "helper" plasmid contains the *vir* genes. Nowadays, many economically important crops are routinely transformed with GOIs using *Agrobacterium*-meditated transformation or the gene gun. A high percentage of the acreage of certain crops is transgenic, such as corn, soybeans, canola, and cotton.

The generation of transgenic plants requires several steps. First, plant expression vectors, that is, binary plasmid containing the gene constructs, are produced. Subsequent steps are required for the manipulation of genes, plasmids, and bacteria, and then finally, plants are transformed. Two types of essential tools—segments of DNA—are required in transformation constructs. (1) An appropriate promoter must be fused upstream of the GOI and also other genes in the construct that are to be transferred into the plant and expressed. These promoters are needed to drive transgene expression. (2) Marker genes are needed to select transgenic plants and/or monitor gene expression.

The choice of promoter is crucial in that it specifies when and where a transgene is expressed in the plant. Expression of marker genes also needs promoters to ensure appropriate expression, which is usually everywhere and all the time in the plant body: constitutive expression.

Since only a very small number of plant cells (e.g., only one in several million to billion cells) can be transformed in most experiments, the selection of those transformed plant cells is a critical step in order to regenerate fully developed transgenic plants from the transformed cells. Marker genes indicate which plant cells have been successfully transformed and aid in their selection. Several kinds of marker genes have been developed and are needed for the diverse roles they play in biotechnology. As the complexity and needs of research increase, there might also arise a requirement to remove marker genes from transgenic plants in order to create marker-free plants.

10.2. PROMOTERS

A promoter is the DNA region that is located upstream of the transcription start site (TSS) of a gene and initiates and controls the transcription of that gene. The nucleotide positions on the promoter region are counted as negative numbers with the first nucleotide upstream to the TSS being −1. The *core promoter* is the minimal promoter region that typically spans from −60 to +40 relative to the TSS (the first transcribed base is at +1) and is capable of properly initiating a transcription (Fig. 10.1). It is estimated that 30–50% of all known core promoters contain a TATA-box located at −45 to −25 bp. The TATA-box (the nucleotide sequence is TATAAA) is the most conserved functional element in promoters and can direct accurate basal transcription initiation. However, some housekeeping and photosynthesis genes do not contain the TATA-box motif; they are referred to as TATA-less promoters. In the TATA-less promoters, the TSSs are controlled by some functionally conserved elements within the core promoters. The sequence 200–300 bp immediately upstream of the core promoter constitutes the *proximal promoter* that contains multiple transcription factor binding sites (i.e., *cis*-regulatory elements or motifs). The distal sequence upstream of the proximal

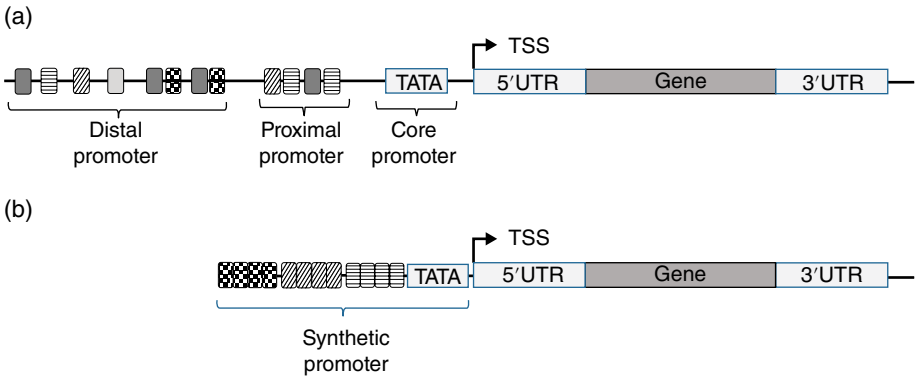


Figure 10.1. Examples of promoters that could be used to regulate the expression of a transgene. In the first example (a) a native plant promoter is shown, which contains a TATA-box in the core region, a few key motifs in the proximal region and many more distant motifs in the distal region. These motifs are the binding sites for various transcription factors, activators, or repressors. In the second example (b) a synthetic promoter has essentially the same features but is in a compressed state. The untranslated regions (UTRs) are shown as is the transcription start site (TSS).

promoter is referred to as the *distal promoter* which may contain additional motifs (i.e., enhancers and repressors) with a weaker influence on the promoter activity.

A promoter sequence needs to be fused upstream of a GOI in the T-DNA for the regulation of gene expression in transformed plant cells. Additionally, a 3′ untranslated region (UTR) must be added to the 3′-end of the GOI for the addition of polyadenylated nucleotides to the 3′-end of the transcribed messenger RNA (mRNA) sequence. In plants, general transcription factors interact with the TATA-box and other key elements in the core promoter to recruit RNA polymerase II to the core promoter in order to form a transcriptional complex and thus initiate transcription of the gene. The binding of the transcriptional complex to the core promoter can bend the promoter back on itself.

Chapter 6 illustrated the importance of promoter sequences in the spatial-temporal regulation of plant genes. Similarly, all transgene constructs, including GOIs and marker genes, require promoters to regulate their transcription reproducibly and predictably. Transgenes can be cloned with a variety of *heterologous* (from another source) promoters for the creation of *chimeric genes* within the binary vectors to provide biotechnologists with a range of expression patterns to suit their needs for targeted trait delivery (see Chapter 9). These promoters may perform a variety of functions, such as constitutive, tissue-specific, and inducible expression of transgenes. In contrast to endogenous plant or viral promoters, which are typically 1000 bp or longer, synthetic promoters can be designed on a computer and synthesized to be very short (<100 bp length) and also confer high transgene expression.

10.2.1. Constitutive Promoters

A constitutive promoter allows constant expression of a gene in all tissues and at all developmental stages. Theoretically, expression from a constitutive promoter comes from the additive effects of interaction of the promoter motifs (short sequence of DNA) with transcription factors (proteins that bind to the DNA sequence) that are constantly present in all cell types, or with different transcription factors present in different cell types. The most frequently used constitutive promoters are those of the 35S transcript in the cauliflower mosaic virus (CaMV) or of genes (i.e., the *nopaline synthase* (*nos*) gene) found in the T-DNA of Ti (tumor-inducing) plasmids. The 35S promoter is generally the stronger promoter of the two and provides an advantage in selection efficiency, particularly in species where the selection procedure is not optimal. Although not of plant origin, both promoters were among the best studied during the time when transformation technologies were first developed

in the 1980s. These two promoters are generally effective over a very wide range of dicot species. Today, many other constitutive promoters have been isolated from highly expressing plant housekeeping genes that function very well in eudicot species, such as the *Arabidopsis thaliana ubiquitin* (UBQ1) promoter and *beta-carotene hydroxylase* promoter, the *Medicago truncatula* MtHP promoter, and the soybean UceS8.3 promoter. However, these promoters were not very effective in monocot species. In contrast, the activity of monocot-derived promoters is always higher in monocots than eudicot promoters. Several monocot gene promoters have been well studied to evaluate their ability for constitutive transgene expression, such as ZmUbi1 from maize (*Zea mays*), OsAct1 and 2, RUBQ1 and 2, and rubi3 from rice (*Oryza sativa*), and PvUbi1 and 2 from switchgrass (*Panicum virgatum*). Among these, the ZmUbi1 and OsAct2 promoters are the most widely used in monocot crops. In contrast to spatiotemporal or inducible promoters (see the text), relatively few promoter elements that are responsible for constitutive expression have been identified. These elements (or domains) are very important. For example, the 35S promoter contains two domains, A and B, which are responsible for specifying transcription. Domain A1 (−90 to −46) confers preferential expression in roots and contains the *as-1* (*activating sequence 1*) element, which is located in the region between −82 and −62 and consists of two CCAAT-box motifs and two TGACG motifs. Domain B (−343 to −90) confers preferential expression in leaf and contains *as-2*, which is located in the −105 to −85 region and composed of motifs GT-1 and Box II. Another good example is the well-conserved G-box-like motif that has been identified in strong constitutive polyubiquitin promoters in both monocots and eudicots, such as in the rice RUBQ2 and 3, maize ZmUbi1, and soybean Gmubi3 and Gmubi7.

Despite such efforts to characterize constitutive promoters for use in plants, there is still a shortage of efficient promoters for high-level constitutive expression, particularly in monocots. Many more constitutive promoters are needed for stacking multiple transgenes in a single transgenic line, which may undergo homology-dependent gene silencing if the same promoter is used repeatedly.

10.2.2. Tissue-Specific Promoters

Spatiotemporal promoters such as tissue- or developmental stage-specific promoters restrict gene expression exclusively to specific tissues or a specific developmental stage. Even though the availability of some well-characterized constitutive promoters provides invaluable sources for gene function analysis and plant genetic engineering, it is also desirable in some instances that transgene expression be temporally and/or spatially regulated. Constitutive transgene expression may be harmful to the host plant, causing increased metabolic burden, abnormal morphology, retarded development, yield penalty, sterility, or transgene silencing. In addition, public concerns regarding the food safety of transgenic plants in many countries have limited the growth of transgenic crops. Using tissue- or developmental stage-specific promoters may avoid transgene expression in seed or fruit and thus relieve this public concern.

During the past years, many different plant promoters have been characterized whose expression is restricted to particular tissues, organs, or developmental stages, such as green tissue-, vascular tissue-, root-, pollen-, and seed-specific promoters (Table 10.1). Well-known green tissue-specific promoters include *Arabidopsis chlorophyll a-/b-binding protein* (Cab3), *Arabidopsis ribulose bisphosphate carboxylase small subunit* (rbcS), and maize *phosphoenolpyruvate carboxylase* (PEPC); all of those are photosynthesis-related genes in plants. The promoters from pumpkin *phloem protein 2* (PP2) and *Arabidopsis profilin 2* (Pfn2) are known for their vascular tissue-specificity, while *Arabidopsis ethylene-insensitive root1* (EIR1) and rice NAC10 are root-specific. Pollen-specific promoters include *late anthogenesis 52* (Lat52) from tomato, tobacco *anther-specific protein 29* (TA29) from tobacco, and maize *pollen specific 13* (Zm13) from maize. The promoters from *Brassica napus napin storage protein A* (*napA*) and rice *glutelin storage protein* (GluB-1) are good representatives for seed-specific promoters.

TABLE 10.1. The Most Widely Used Tissue-Specific Promoters in Plants

Promoter type	Promoter name	Gene function	Species	References
Green tissue	Cab3	Chlorophyll a-/b-binding protein	*Arabidopsis*	Mitra et al. (1989)
	rbcS	Ribulose bisphosphate carboxylase small subunit	*Arabidopsis*	De Almeida et al. (1989)
	PEPC	Phosphoenolpyruvate carboxylase	Maize	Ku et al. (1999)
Vascular tissue	PP2	Phloem protein 2	Pumpkin	Guo et al. (2004)
	Pfn2	Profilin 2	*Arabidopsis*	Christensen et al. (1996)
Root	EIR1	Ethylene-insensitive root1	*Arabidopsis*	Luschnig et al. (1998)
	NAC10	*N*AM, *A*TAF1-2, *C*UC2	Rice	Jeong et al. (2000)
Pollen	Lat52;59	Late anthogenesis	Tomato	Twell et al. (1991)
	TA29	Tobacco anther-specific protein TA29	Tobacco	Koltunow et al. (1990)
	Zm13	Pollen specific	Maize	Hamilton et al. (1998)
Seed	napA	Napin storage protein	*Brassica napus*	Rask et al. (1998)
	GluB-1	Glutelin storage protein	Rice	Wu et al. (2000)

The specificity of tissue-specific promoters is determined by interactions between tissue-specific transcription factors and their respective DNA binding sites or motifs within the tissue-specific promoters. Even though most of these tissue-specific transcription factors are still unknown, some of the motifs responsible for tissue-specific expression have been identified within tissue-specific promoters. For example, the CANBNNAP motif activates endosperm- and leaf-specific gene expression. The RGATAOS motif promotes phloem-specific gene expression, while motifs ACGTROOT1, ROOTMOTIFTAPOX1, WUSATAg, OSE1ROOTNODULE, and OSE2ROOTNODULE stimulate root-specific gene expression. The POLLEN1LELAT52 element functions for pollen-specific expression. RYREPEATBNNAPA and TGTCACACMCUCUMISIN are responsible for seed- and fruit-specific gene expression, respectively. These known tissue-specific motifs can be used for the engineering of synthetic promoters for use in plants (see the text).

10.2.3. Inducible Promoters

Inducible promoters turn on gene expression in response to exposure to specific inducers. In the absence of an inducer, a gene driven by an inducible promoter will not be transcribed. The inducer directly or indirectly converts the factors that specifically bind to an inducible promoter and activate gene transcription from an inactive state to an active state. Thus, gene expression under the control of inducible promoters can be artificially controlled using inducers. An ideal inducible expression system should have the following features (Corrado and Karali 2009): (a) expression of inducible promoters in the absence of the signals should be extremely low, leading to very little background expression; (b) inducible promoters should quickly and significantly respond to their specific signals only in a dose-dependent manner; (c) external signals should not have any pleiotropic effects on endogenous gene expression or be toxic to plants; (d) external signals should be easy to control—say, by spraying an innocuous chemical that acts as a signal. With these features, an ideal inducible expression system should allow gene expression to be switched on or off by adding or removing specific signals. Such external signals can be biotic factors (i.e., pathogens and disease agents), abiotic factors (i.e., light, heat, cold, salt, and wounding), and chemicals (i.e., proteins, metabolites (sugar, and alcohol), growth regulators, herbicides, or phenolic compounds).

An important type of inducible promoters are those that turn on gene expression in plants upon pathogen infection (Table 10.2). Good examples here are promoters from *Arabidopsis thaliana*

TABLE 10.2. The Most Widely-used Inducible Promoters in Plants

Promoter inducibility	Promoter name	Gene function	Species	References
Pathogen	PR1	Pathogenesis-related 1	*Arabidopsis*	Lebel et al. (1998)
	NPR1	Nonexpressor of PR1	*Arabidopsis*	Yu et al. (2001)
	VSP1	Vegetative storage protein 1	*Arabidopsis*	Guerineau et al. (2003)
	PcPR1-1	Pathogenesis-related 1	Parsley	Rushton et al. (2002)
	PcPAL1	Phenylalanine ammonia-lyase 1	Parsley	Lois et al. (1989)
	PR2-d	Pathogenesis-related 2-d	Tobacco	Shah et al. (1996)
	NtGlnP	Glucanase 2	Tobacco	
Light	CHS	Chalcone synthase	Parsley	Weisshaar et al. (1991)
	LHCP	Light-harvesting chlorophyll a/b protein	Pea	Simpson et al. (1985)
	Rca	Rubisco activase	Spinach	Orozco and Ogren (1993)
Wound	MPI	Maize proteinase inhibitor	Maize	Cordero et al. (1994)
	Pin2	Proteinase inhibitor II	Potato	Thornburg et al. (1987)
Drought	ERD1	Early responsive to dehydration stress 1	*Arabidopsis*	Tran et al. (2004)
Salt	RD29A, B	Responsive to desiccation 29A, B	*Arabidopsis*	Yamaguchi-Shinozaki and Shinozaki (1994)
Cold	Cor15A	Cold-regulated 15A	*Arabidopsis*	Stockinger et al. (1997)
	CBF2/DREB1C	C-repeat/DRE binding factor 1	*Arabidopsis*	Zarka et al. (2003)
ABA	HVA22	ABA-inducible	Wheat	Shen et al. (1993)
	Osem	Rice homolog of Em	Rice	Hattori et al. (1995)
ABA, Drought	RD22	Responsive to desiccation	*Arabidopsis*	Abe et al. (1997)
	Em	Late embryogenesis	Wheat	Guiltinan et al. (1990)
	RD29A,B	Responsive to desiccation 29A, B	*Arabidopsis*	Yamaguchi-Shinozaki and Shinozaki (1994)
Ethanol	AlcA	Alcohol-regulated	*Aspergillus nidulans*	Caddick et al. (1998)

PR1, NPR1, and VSP1; parsley PcPR1-1 and PcPAL1; *Phaseolus vulgaris chalcone synthase* (PvCHS15); and tobacco PR2-d, chitinase, glucanase 2 (NtGlnP), and NtPR1-a. A drawback of some pathogen-inducible promoters is that they may react only to certain types of pathogens. The most extensively studied light-inducible promoters are parsley CHS, pea *light-harvesting chlorophyll a/b protein* (LHCP), and spinach *rubisco activase* (Rca). Maize *proteinase inhibitor* (MPI) and potato *proteinase inhibitor II* (Pin2) are good representatives of wound-inducible promoters. Many plant promoters are induced by drought, high salinity, cold stress, and/or abscisic acid (ABA). These promoters include those from *Arabidopsis early responsive to dehydration stress 1* (ERD1), *responsive to desiccation 22* and 29A,B (RD22; RD29A,B), *cold-regulated 15A* (Cor15A), and *C-repeat/DRE binding factor 1* (CBF2/DREB1C), from wheat *ABA-inducible 22* (HVA22) and *late embryogenesis* (Em), and from rice *homolog of Em* (*Osem*). These stress-inducible promoters are responsive to one or several environmental stresses (i.e., drought, salt, cold, or ABA). Since some of the biotic and abiotic factors are difficult to control outside of experimental settings, chemical inducible promoters, such as antibiotics, alcohol, steroids, herbicides, and heavy metals, are of particular interest. A well-studied alcohol-inducible promoter in plants is the *alcohol-related* (AlcR) promoter from fungi *Aspergillus nidulans*.

It is the specific binding of transcription factors to the motifs within inducible promoters that confers inducible gene expression in response to different inducers. Many motifs that function in

inducible promoters have been precisely characterized to elucidate the molecular mechanisms of inducible gene expression. Pathogen- or elicitor-responsive elements include motifs PR1, W-box, GCC, Gst1, D-box, S-box, ERE, SARE, JAR, NPR1-motif, and so on. The LRF-1 motif is responsible for light-inducible expression in tomato *rbcS* genes. The G box element with the core sequence CAGTG is essential for transcriptional activity in the *Arabidopsis rbcS*, wheat *Em*, and parsley *CHS* genes. The C-repeat (CRT) and dehydration-responsive element (DRE) are responsible for drought-, high salt-, and cold-inducible gene expression. The low-temperature-responsive element (LTRE) functions in cold-inducible gene expression. The ABA-responsive element (ABRE) PyACGTGGC works for ABA-inducible expression.

10.2.4. Synthetic Promoters

The promoter regions of plant genes have modular functional motifs being embedded in hundreds to thousands of nonfunctional DNA nucleotides; therefore, endogenous plant promoters are always very long, weak, and complex. One strategy to overcome this complexity is to fuse motifs of interest together in order to make synthetic functional promoters, thereby reducing expression profile complexity while increasing promoter strength. Synthetic promoters are generated by fusing together promoter motifs from sometimes diverse origins upstream of a core promoter region (often the TATA-box region; Fig. 10.1b). The use of motifs can be fine-tuned to better allow for the regulation of spatiotemporal gene activity and multigene expression, and to overcome drawbacks such as homology-dependent gene silencing.

Synthetic promoter engineering in terms of motif composition and promoter architecture requires accurate deciphering of motifs found in endogenous plant promoters. The most commonly used core promoter is the minimal 35S (i.e., −46 35S) promoter. Motifs fused together with any core promoters can be selected from sequences that have been characterized experimentally from known plant motif databases, or from *de novo* motif discovery (see Chapter 17). Synthetic promoters can be engineered for functions of constitutive, bidirectional, or inducible expression of GOIs (Table 10.3). For example, the well-known *tet* operator can be fused together with the 35S promoter to make a synthetic promoter, Triple-Op, for conditional repression, in which the TetR repressor specifically binds to the *tet* operator and inhibits gene expression driven by the promoter containing the operator. Upon association with tetracycline, the repressor is released from the promoter, resulting in a de-repression effect on gene expression. Another good example is the *LexA* operator-XVE-inducible system. XVE is a fusion protein containing the DNA binding domain of the bacterial repressor LexA, the VP16 activation domain and the regulatory region of the rat glucocorticoid receptor (GR), and it specifically binds to the LexA operator. Upon addition of β-estradiol, XVE can activate gene expression driven by the LexA operator promoter fused to the minimal 35S promoter.

10.3. MARKER GENES

Marker genes are used in plant transformation experiments to determine if a transgene has been successfully transformed into a plant genome or to select for transgenic tissues and plants among nontransgenics. In most cases, the marker genes, which are selected or observed, are linked together with the transgene of interest. Therefore, it is used as an indirect selection of transgenic plant (Fig. 10.2). Various marker genes have played crucial roles in facilitating the production of transgenic plants, the subsequent identification of the transgenic plants, and the fine-tuning of procedures that are needed to increase transformation frequencies. Marker genes fall into two categories: *selectable* marker genes and *reporter* genes (also referred to as *screenable* marker genes or *visual* marker genes) (see Table 10.4 for examples).

TABLE 10.3. Synthetic Promoters that have been Designed for the Regulation of Transgene Expression in Plants

Promoter	Pattern of expression	Source(s)	Species tested	References
Mod2A1T, Mod3A1T	Constitutive	CaMV 35S	Tobacco	Bhullar et al. (2003)
Amas'AocsPocs, (Aocs)3AmasPmas	Constitutive	Mannopine synthase and octopine synthase promoters	Tobacco	Ni et al. (1995)
Pcec, Pmec	Constitutive	Transcription activation module (TAM) designed from a database of highly expressed plant genes	Tobacco, tomato	Sawant et al. (2001)
SynPro3, SynPro5	Constitutive, stronger in leaf	Synthetic directly repeated elements, CaMV 35S	Tobacco	Cazzonelli and Velten (2008)
Mac	Constitutive, wound-inducible	CaMV 35S, Ti plasmid mannopine synthetase promoter	Tobacco, tomato	Comai et al. (1990)
pd35GR, pd35ER, pdCGR, pq35GR	Bidirectional, constitutive	CaMV 35S or CsVMV	Grape, tobacco	Li et al. (2004)
P1301A; P1301B	Bidirectional salicylic acid-, salinity-, and IAA-inducible	Transcription activation module (TAM) designed from a database of highly expressed plant genes	Tobacco	Chaturvedi et al. (2006)
pGLbd1-6	Bidirectional, constitutive or wounding-, JA-, and leaf senescence-inducible	CaMV 35S, *PClSV*, or *OPR1* promoters	*Arabidopsis*	Xie et al. (2001)
4×GT1,4×GATA 4×G, 2×Z	Light-inducible	*Arabidopsis* light-regulated photosynthetic genes (cab, *rbcS*) promoters	*Arabidopsis*	Puente et al. (1996)
Triple-Op	Tetracycline-inducible	CaMV 35S, *tet* operator	Tobacco	Gatz et al. (1992)
CGEcR, VGEcR	Methoxyfenozide-inducible	European corn borer ecdysone receptor	Maize	Unger et al. (2002)
GRE	Dexamethasone-inducible	GAL4 UPS, CaMV 35S	Tobacco *Arabidopsis*	Aoyama and Chua (1997)
4×ERE	β-estradiol-inducible	Estrogen-responsive element	Maize	Bruce et al. (2000)
O_{LexA}-46	β-estradiol-inducible	LexA operator sequence, minimal 35S	*Arabidopsis* Tobacco	Zuo et al. (2000)
palcA	Ethanol	*Aspergillus nidulans* alcA promoter, minimal 35S	Tobacco	Caddick et al. (1998)
TCS	Cytokinin-inducible	B-type *Arabidopsis* response regulator (ARR)-binding motifs and minimal 35S	*Arabidopsis*	Muller and Sheen (2008)
DR5(8×)	Auxin-inducible	Soybean *GH3* promoter	*Arabidopsis*, Carrot	Ulmasov et al. (1997)
4×W1, 4×W2, 4×D, 4×GCC, 4×S, 4×JERE, 4×GST, 4×DRE	Pathogen-and would-inducible	Parsley *PR1*, Tobacco chitinase, Parsley *ELI7* Periwinkle *Str*, Potato *GST1* and *Arabidopsis rd29A* gene promoters	*Arabidopsis*	Rushton et al. (2002)
4H1-46	Seed-specific and desiccation-, salinity-, and abscisic acid-inducible	Tobacco *hex-1* promoter	Tobacco	Lam et al. (1991)

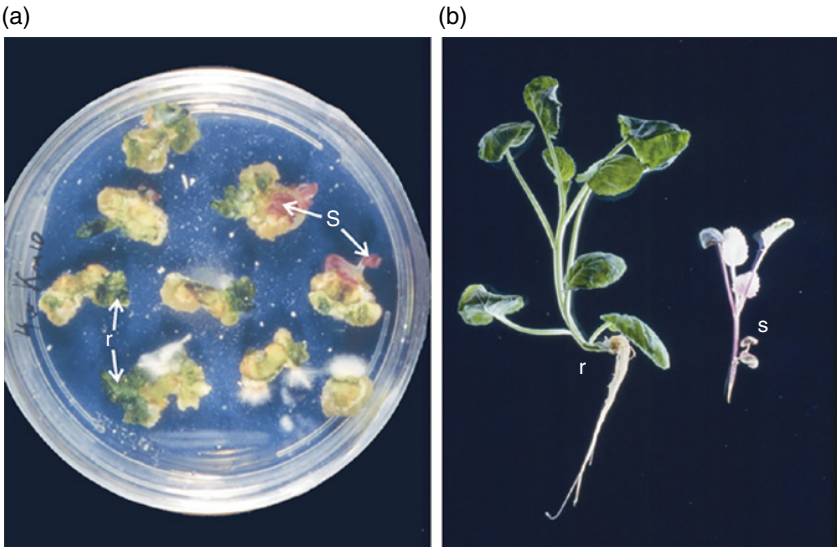


Figure 10.2. Selection of transgenic canola (*Brassica napus* cv Westar) on kanamycin-containing tissue culture media. Stem explants were first infected with an *A. tumefaciens* strain harboring a transformation vector with a chimeric *nptII* gene designed to confer kanamycin resistance on transformed plant tissue. (a) After cocultivation of plant tissue with *Agrobacterium* allowing transformation to occur, the plant tissues were transferred to tissue culture media containing kanamycin for growth of callus tissue and shoot differentiation. Much of the non-transformed tissues turned white (see arrows pointing to "s") and stopped growing because they were sensitive to the antibiotic. Transformed tissues remained green and continued to grow and differentiate because they were resistant to kanamycin (see the arrows pointing to "r"). (b) Transgenic shoots that differentiated in the presence of kanamycin were excised from the callus and transferred to media for the regeneration of roots. Escapes that were not truly kanamycin-resistant were unable to regenerate roots in the presence of the antibiotic. (*Source*: Courtesy of Pierre Charest). (*See insert for color representation of the figure.*)

TABLE 10.4. Categories of Marker Genes and Selective Agents used in Plants

Category	Marker genes	Source of genes	Selective agent
Selectable marker genes:			
Antibiotic-resistant	*nptII*, *neo*, *aphII*	*Escherichia coli* Tn5 (bacterial)	Kanamycin
	hpt, *hph*, *aphIV*	*E. coli* (bacterial)	Hygromycin
Herbicide-resistant	*bar*	*Streptomyces hygroscopicus* (bacterial)	Phosphinothricin
	pat	*Streptomyces viridochromogenes* (bacterial)	Phosphinothricin
	CP4 EPSPS	*Agrobacterium* sp. strain CP4 (bacterial)	Glyphosate
Nutritional inhibitor-related	*manA*	*E. coli* (bacterial)	Mannose
	xylA	*S. rubiginosus*; *Thermoanaerobacterium thermosulfurogenes* (bacterial)	D-xylose
Hormone-related	*ipt*	*Agrobacterium tumefaciens* (bacterial)	N/A[a]
Ablation	*codA*	*E. coli* (bacterial)	5-Fluorocytosine
Reporter genes:			
Enzymatic	*uidA*, *gusA*	*E. coli* (bacterial)	MUG, X-gluc
	Luc	various	luciferin
Fluorescent proteins	*gfp*	*Aequorea victoria* (jellyfish)	N/A
	pporRFP	*Porites porites* (hard coral)	N/A
	mOrange	*Discosoma* sp. (soft coral)	N/A

[a] Not applicable.

10.3.1. Selectable Marker Genes

A selectable marker gene is useful in plant transformation and regeneration by conferring a selective growth advantage to plant cells so that transformed cells can outgrow the non-transformed cells, thus being selected. Because only a few cells are transformed in a population of target cells, there would be little chance of recovering transgenic cells without selectable markers. Most selectable marker genes can protect the transformed plant cells—through the action of detoxification or resistance mechanisms encoded by the selectable marker genes—from a selective agent in the plant growth media that would normally kill non-transgenic tissues or inhibit their growth. In addition, a few selectable marker genes do not require the use of selection agents, yet they promote the selective growth and differentiation of transformed tissues. An example is the *ipt* gene that enhances transformed shoot development by mediating plant hormone levels (see further).

Over 50 selectable marker genes have been described in the literature, primarily for nuclear transformation. The underlying principles used to achieve selection differ widely among the selectable marker genes, and the terminology for describing them in the literature has been confusing. Table 10.4 provides a classification system for the various marker genes used in plants, such as antibiotic-resistant, herbicide-resistant, nutritional inhibitor-related, hormone-related, and ablation genes. A more comprehensive list can be found in Miki and McHugh (2004). Having many different systems is important as they vary in efficiency among plant species. Furthermore, experiments are often required in which different transgene insertions are combined in individual plant lines through genetic crosses using separate parental transgenic lines or through consecutive transformation steps. Different selectable marker genes allow the researcher to follow the segregation of each insertion event independently. However, the scientific literature shows that only a few selection systems have been adopted routinely to generate transgenic plants for research or for commercialization. These include the *nptII* and *hpt* genes, which confer resistance to the antibiotics kanamycin and hygromycin, respectively, and the *bar* or *pat* genes, which confer resistance to the herbicide phosphinothricin (Table 10.4). In field trials, the most frequently used selectable marker genes are the *nptII* and *bar/pat* genes (Miki and McHugh 2004).

The use of selectable maker genes is very important because the ease with which transformed cells are allowed to proliferate must be balanced with the stringency with which the non-transformed cells are suppressed or killed. The accumulation of toxins from dead tissues can adversely affect the ability of living tissues to survive, particularly if they are present in limited numbers within a larger population of dying or dead materials. The optimal selection conditions tend to be specific for each plant species and tissue type. If not properly administered, the proportion of transgenic materials may be disproportionately low relative to the frequency of transformation events that actually occurred. Conversely, if the frequency of "escapes" (i.e., non-transgenic tissues that the researcher believes to be transgenic since they survived selection) is too high, then considerable effort and cost would be needed to separate the transgenic lines from the non-transformed material later. To be effective, the selectable marker genes should not interact with specific targets within the plant or alter signal transduction pathways in a way that changes the plant. If they create such changes it would be difficult to identify the phenotypes associated with the GOI or the factors affecting their expression.

Once transgenic plants have been recovered, the selectable marker gene can act as a genetic marker for subsequent genetic studies as it is linked to the GOI (Fig. 10.3). For example, the selectable marker gene allows the researcher to predict the number of segregating insertion events that have occurred in a transgenic line and also to monitor the transmission of the linked transgenes among the progeny of the plant. The genetic analysis of marker gene segregation is usually an important step in selecting the homozygous transgenic lines with single insertions (i.e., with a simple 3 : 1 Mendelian segregation) used for detailed studies of the GOI.

Most selectable marker genes are of bacterial origin (Table 10.4). Some of these are altered to be efficiently expressed in plant cells because the regulatory signals on the bacterial genes will not be

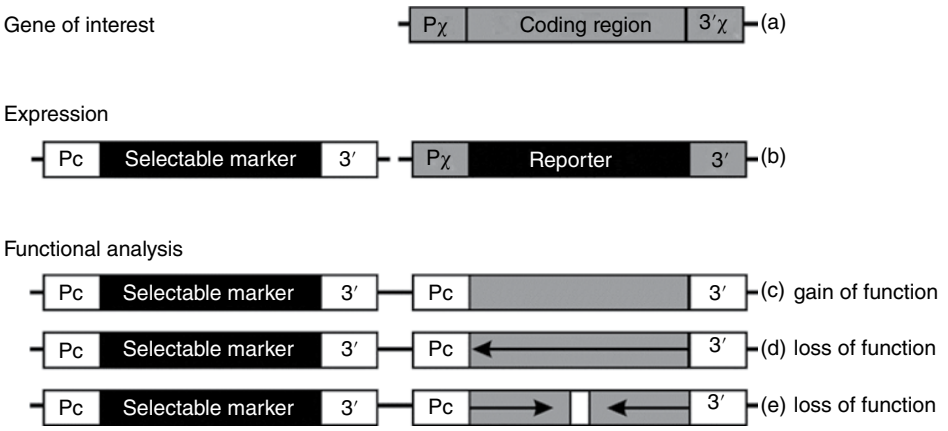


Figure 10.3. Functional organization of selectable marker genes and reporter genes on transformation vectors used to transfer DNA to plant cells. The selectable marker genes are a fundamental component of the transformation vectors as they are needed for the recovery of transgenic material. The vectors are used for many purposes, including the study of plant genes and their regulatory elements. Often the function of genes emerging from genomics studies are unknown and the transgenic plant provides an experimental model for gaining functional understanding. (a) The gene of interest (GOI) can be examined in many ways. The regulatory elements are often found in the noncoding regions of the gene. For example, the promoter (Pχ) is found in the 5′ upstream region and includes a number elements needed for transcription, including the core promoter and often enhancer or repressor elements. Some of these elements may also exist in the 3′ end region. (b) By fusing the 5′ and 3′ noncoding regions to a reporter gene and inclusion of the chimeric gene in the transformation vector, the patterns of gene regulation can be assessed in transgenic plants. (c) Gain-of-function experiments can be performed by the overexpression of the coding region using a strong constitutive promoters (Pc), such as the 35S promoter, and 3′ ends needed for termination and polyadenylation (3′), such as those from the *nos* gene or 35S transcript. A phenotype in the transgenic plant may reveal function. (d) A mutant phenotype may also be mimicked by eliminating or reducing the expression of the GOI by creating an antisense transcript in the transgenic plant. (e) This may also be achieved by creating a vector with inverted repeats of the gene of interest, which may induce gene silencing. In each case, the selectable marker and the reporter genes serve different purposes.

correctly recognized by the plant gene expression machinery. The modifications could include changes to the codons favored by plants and elimination of cryptic sites that could result in aberrant processing of transcripts. They could also include swapping of the upstream and downstream regulatory elements with plant sequences to create chimeric genes that will be recognized by the plant transcriptional and translational systems.

10.3.1.1. Antibiotic Resistance Genes. Many antibiotics used as selective agents in the plant growth media are aminoglycoside antibiotics. These antibiotics contain amino-modified glycosides (sugar) or any aminosugar substructures; they inhibit protein synthesis in Gram-negative bacteria as well as in eukaryote plastids and mitochondria by irreversibly binding to the ribosomal subunits. Good examples are kanamycin, hygromycin, and gentamicin derivative G418 that are all very toxic to plant, animal, and fungal cells. Kanamycin is produced as a trisaccharide composed of a deoxystreptamine and two glucosamines found in soil actinomycete *Streptomyces kanamyceticus*. Hygromycin B is an aminocyclitol antibiotic produced in the bacterium *Streptomyces hygroscopicus* and has broad spectrum activity against bacteria, fungi, and eukaryotes. Gentamicin is synthesized in a genus of the Gram-positive bacteria, *Micromonospora*, and is composed of a mixture of related gentamicin components that are antagonist to Gram-negative bacteria.

All of the antibiotics-resistant selectable marker genes have been derived from bacterial sources. These selectable marker genes encode enzymes that detoxify aminoglycoside antibiotics via

modifications to the molecular structures of the antibiotics, leading to the prevention of the binding of antibiotics to the ribosomal subunits, and thus antibiotics resistance in the transformed cells. The modifications of the aminoglycosides in the antibiotics include ATP-dependent O-phosphorylation by phosphotransferases, acetyl CoA-dependent *N*-acetylation by acetyltransferases, and ATP-dependent O-adenylation by nucleotidyl-transferases. Here, we only discuss two phosphotransferase genes.

Neomycin phosphotransferase II (NPTII) from *Escherichia coli* transposon Tn5 is a bacterial aminoglycoside 3-phosphotransferase II (APH[3]II; EC2.7.1.95). NPTII catalyzes the ATP-dependent phosphorylation of the 3-hydroxyl group of the amino-hexose portion of certain aminoglycoside antibiotics such as kanamycin and gentamicin (G418). The gene coding for NPTII (*nptII*, also designated *neo* or *aphII*) was the first selectable marker gene to be tested in plants, created by fusing to the 5′-and 3′-end regulatory elements from the *nopaline synthase* gene (*nos*) from the T-DNA of the *A. tumefaciens* Ti plasmid. The *nos* gene elements in the synthesized chimeric gene confer constitutive expression of the *nptII* gene and thus kanamycin resistance within all cells of the transgenic plant. A stronger upstream promoter sequence from 35S RNA generates a higher level of *nptII* gene expression, which results in a higher level of kanamycin resistance. The *nptII* gene can function as a selectable marker in both the nuclear and plastid genomes. So far, the *nptII* gene has become the most widely used selectable marker gene in plants. The popularity of kanamycin resistance, conferred by the *nptII* gene, is because kanamycin is very effective, functions in a wide range of plant species, and appears to be very safe for use in food and feed crops. Since it also functions effectively in a wide range of microorganisms and eukaryotic cells, some initial concerns had been expressed about the potential transfer of antibiotic resistance to other organisms. However, it has been used since the mid-1980s in crops, and no adverse effects on humans, animals, or the environment have yet to appear (Flavell et al. 1992). It is also known that expression of the *nptII* gene in plants does not alter the patterns of transcription in plants, so that transgenic plants expressing it are essentially equivalent in composition to nontransgenic plants.

The *E. coli hph* (also known as *hpt*, *aphIV*) gene codes for hygromycin phosphotransferase (HPT, EC 2.7.1.119), and confers hygromycin B resistance in bacteria, fungi, animal cells, and plant cells by detoxifying hygromycin B via an ATP-dependent phosphorylation of a 7-hydroxyl group. This gene has been used as a selectable marker when *nptII* was not found to be effective. Hygromycin B has become the second most frequently used selectable marker gene for selection. However, it is very toxic to plant cells relative to kanamycin and more difficult to apply without "overkill."

10.3.1.2. Herbicide-Resistant Genes. Herbicides are used to kill unwanted weeds while leaving the desired crops relatively unharmed or less harmed. A well-known example is the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D), a synthetic auxin that selectively kills some broadleaf weeds in cereals without harming the crops. Phosphinothricin (PPT), the ammonium salt of glufosinate, is the active component of several commercial herbicides such as Basta, Buster, Biolophos, and Liberty. As an analog of L-glutamic acid, it is a competitive inhibitor of glutamine synthase (GS) that is essential for the assimilation of ammonia into plants. By inhibition of GS, ammonia accumulates to toxic levels and kills weeds. Glyphosate (*N*-(phosphonomethyl)glycine) is another commercial herbicide under the trade name "Roundup" that kills weeds, especially annual broadleaf weeds and grasses. It inhibits 5-enolpyruvoyl-shikimate-3-phosphate synthetase (EPSPS), an enzyme involved in the synthesis of aromatic amino acids tyrosine, tryptophan, and phenylalanine.

The enzyme phosphinothricin *N*-acetyltransferase (PAT) detoxifies PPT by acetyl CoA-mediated acetylation. The acetylated form of PPT is unable to bind to GS, leading to Basta resistance. Two genes coding for this enzyme have been cloned: the *bar* gene for *bia*lophos *r*esistance (bialaphos consists of two L-alanine residues and PPT) from *S. hygroscopicus* and the *pat* gene (*p*hosphinothricin *a*cetyl*t*ransferase) from *S. viridochromogenes*. Both have been extensively used as selectable marker genes, particularly among cereal species where kanamycin selection may be less efficient. Typically, kanamycin does not kill monocots very effectively, whereas bialophos or PPT does.

Plants containing the *bar* or *pat* genes have been among the first to receive regulatory approval for unconfined field production and have been assessed as safe by a number of international regulatory agencies.

A version of the *EPSPS* gene was found to be resistant to glyphosate inhibition in some microorganisms. For example, the *CP4 EPSPS* gene from *Agrobacterium* sp. strain CP4 was found to contain a single mutation (Gly96 to Ala100) which confers glyphosate resistance, and has been transformed into plants for commercial purposes. Currently, commercial glyphosate-resistant crops include soybean, maize, sorghum, canola, alfalfa, and cotton.

10.3.1.3. Nutritional Inhibitor-Related Genes. Most conditional positive selection systems use toxic substrates (i.e., antibiotics or herbicides) for selection of the transformed tissues; however, the use of nontoxic metabolic intermediates has emerged as an alternative. This type of system differs from the use of antibiotics, herbicides, or drugs in that nutritional inhibitors serving as carbon sources are restricted from use by the plant cells unless provided with an enzyme that allows entry of the carbon source into primary metabolism. Examples of such selective agents include mannose and D-xylose. Bacterial genes *manA* from *E. coli* and *xylA* from *S. rubiginosus* and *Thermoanaerobacterium thermosulfurogenes* encode for phosphomannose isomerase (PMI; EC 5.3.1.8) and xylose isomerase (XYI; EC 5.3.1.5), respectively, and provide the enzymes that allow entry into glycolysis (Fig. 10.4). These genes are present in microorganisms but absent in many plant species. The *manA* gene functions in the conversion of mannose-6-phosphate to fructose-6-phosphate, an intermediate of glycolysis, and confers on transformed plant cells the ability to use mannose as a carbon source. The *xylA* gene converts xylose to fructose-6-phosphate and allows transformed plant cells to use D-xylose as a carbon source. The apparent advantage is that these resistance genes work with a wide range of plant species and appear to yield higher transformation frequencies because the selection is not as harsh as with toxic chemicals such as antibiotics. This approach differs fundamentally from the others discussed so far in that the novel trait encoded by the selectable marker gene alters a basic aspect of plant metabolism.

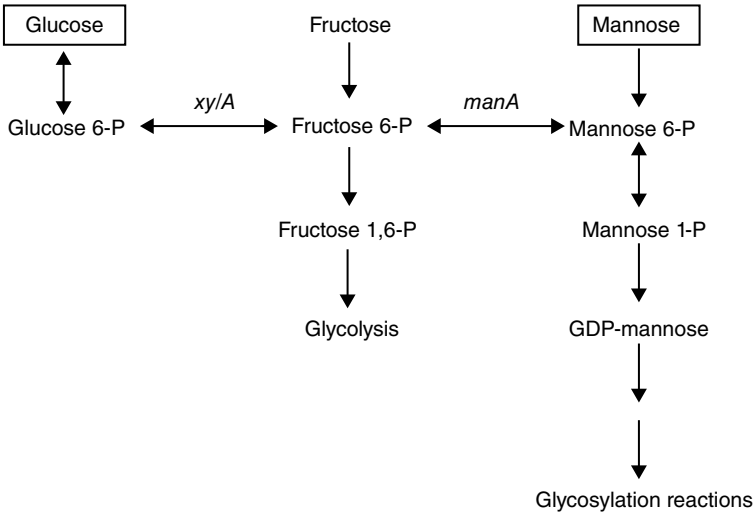


Figure 10.4. Basic metabolism involving D-xylose and mannose. Xylose is a monosaccharide of aldopentose that contains five carbon atoms and a formyl functional group. Xylose isomerase encoded by the *xylA* gene function in the reversible conversion of glucose 6-phosphate (P) to fructose 6-P. Phosphomannose isomerase encoded by the *manA* gene reversibly converts mannose 6-P to fructose 6-P. With either D-xylose or mannose as the carbon source in the selection media, transformed plant cells with the selectable marker genes *xylA* or *manA* are able to grow, while non-transformed plant cells will starve to death with the absence of either gene.

10.3.1.4. Hormone-Related Marker Genes. A few selection systems do not require any substrates (agents) for selection. This hormone-related selection system is based on the use of genes that confer a growth advantage, distinguishable morphology, or that selectively induce the differentiation of transformed tissues but do not necessarily kill non-transgenic tissues. The use of shoot organogenesis to select for transformed tissues is the most advanced example. Shoot formation in culture depends on the presence of high cytokinin: auxin ratios. The T-DNA of the Ti plasmids from *A. tumefaciens* contains an isopentenyl transferase (*ipt*) gene encoding the enzyme IPT, which catalyzes the first step in cytokinin biosynthesis, that is, the synthesis of isopentenyl adenosine-5′-monophosphate (Table 10.4). Expression of the *ipt* gene alone in plant cells results in a higher frequency of shoot regeneration and thus recovery of transformed material. The difficulty is that the shoots have abnormal morphology due to the cytokinin imbalance and cannot produce roots (Ebinuma et al. 2001). To overcome this obstacle, an inducible promoter, such as the β-estradiol-inducible promoter, is needed to restrict the timing of expression of the *ipt* gene (Zuo et al. 2001). A number of alternatives have been demonstrated to have potential; however, these need time to be fully evaluated and developed. Again, this approach differs from most other systems in that it intervenes in the basic processes of plant cell growth and differentiation.

10.3.1.5. Ablation Genes. The ablation class of selectable marker genes can play an important role in experiments by eliminating unwanted transformation events or when selecting against expression in specific tissues or under specific inducible conditions. One example is the *E. coli codA* gene, which encodes for cytosine deaminase (Table 10.4). Cytosine deaminase mediates the uptake and conversion of cytosine into uracil, and is absent in plants. The *codA* gene can be used as a selectable marker gene in the presence of cytosine. Transformed plant cells containing the bacterial *codA* gene can survive in the presence of cytosine, whereas non-transformed cells cannot. It is interesting that this gene has been shown to be effective in nuclear and plastid transformation. In addition, this class of selectable marker genes can also be used to kill transformed plant cells or specific plant tissues when driven by tissue-specific promoters or inducible promoters.

Another example is the *barnase* gene, which has been used to create male-sterile plants. As shown in Figure 10.5, the use of the *barnase* gene expressed only in tapetum cells, can ablate these cells, and therefore no viable pollen is produced. The consequence to the plant is the inability to produce pollen or male sterility (Mariani et al. 1990). The activity of *barnase* can be controlled by a specific protein inhibitor, *barstar*, which is also found in the same bacterium.

10.3.2. Reporter Genes

Whereas selectable marker genes help the researcher select transgenic tissues, reporter genes usually report on the activity of promoters of interest, and/or which cells are transgenic. For promoter activity, the researcher fuses the promoter to the reporter gene and subsequently observes marker gene activity. They can also be used for gene expression assays when fused to a GOI to create a gene fusion. Thus, a reporter gene is a gene that is attached to any promoter or GOI so that the transformed plant cells expressing the reporter can be easily identified. It is used as an indicator of transgene expression in plants.

As important as the selectable marker genes have been for the development of transformation technologies, reporter genes have played a different fundamental role in the growth of our understanding of gene regulation mechanisms in plants. These kinds of genes have formed an invaluable partnership with selectable marker genes in transgenic research. Commonly used reporter genes often exhibit visually identifiable characteristics such as fluorescence or luminescence that result in changes of the appearance of plant tissues. Thus, reporter genes are sometimes called “visible marker genes.” Although several reporter genes have been described, three have been particularly influential and have dominated the scientific literature. These are the genes coding for β-glucuronidase (GUS),

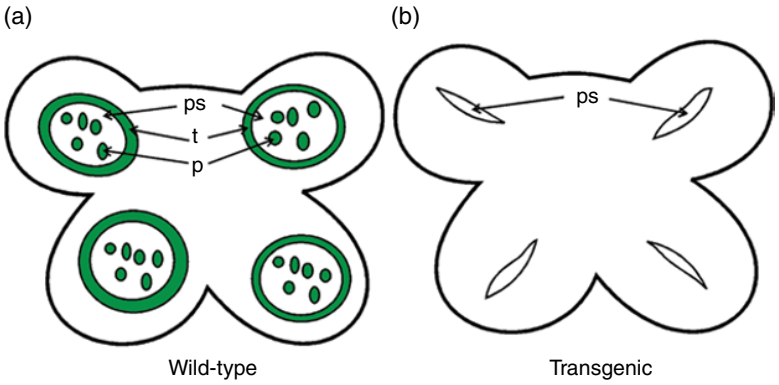


Figure 10.5. Diagrammatic illustration on the use of *barnase* as a negative selectable marker gene for the ablation of the canola tapetal cells: (a) wild-type and (b) transgenic. Barnase codes for a ribonuclease from *Bacillus amyloliquefaciens*. When it is expressed in plants under the control of the tapetum-specific promoter (TA29), expression was restricted to the cells of the tapetum (t) in transgenic plants. The ribonuclease activity in the tapetum resulted in failure of the tapetum to develop and collapse of the pollen sac (ps). Because the tapetal cells are the precursors of the pollen cells, pollen (p) cannot differentiate in the transgenic plants and the plants are therefore male sterile. Because the pattern of ribonuclease expression was tapetum-specific, the rest of the plant was unaltered. (*Source*: From Mariani et al. (1990).)

luciferase (LUC), and fluorescent proteins (FPs). GUS and LUC require the use of an external substrate for detection of activity, whereas FPs encode a protein that is directly detectable without the use of a substrate.

10.3.2.1. β-Glucuronidase. The *E. coli uidA* gene (also known as *gusA* or simply, *gus*) encoding β-glucuronidase (GUS) has been the most widely used reporter system in plants. It allows the analysis of promoter activities in a quantitative or qualitative way by converting its specific colorless (5-bromo-4-chloro-3-indolyl glucuronide; X-gluc) or non-fluorescent (4-methylumbelliferyl glucuronide; MUG) substrates into colored (blue) or fluorescent products, respectively (Fig. 10.6). The staining pattern using X-gluc as its substrate indicates the strength and patterns of the promoters driving the *uidA* gene. The quantitative measurement of the enzyme using MUG as its substrate reflects the relative expression level of the *uidA* gene, which in turn indicates the strength of the promoters fused to the reporter gene. Thus, plant researchers are able to tell which plants are transgenic, and how the promoters work in plant tissues. The enzyme is stable in plant cells and can accumulate to high levels without toxicity to the plant cells. It confers no apparent phenotype to plants in the absence of its substrates and therefore can be used to study plant processes without concern of artifacts resulting from non-target or pleiotropic effects. It has been used in transcription fusions to study a wide range of regulatory elements cloned from the plant genome (Fig. 10.3) and also for promoter-trapping experiments (Fig. 10.7). It also forms stable translational fusions with proteins; for example, fusions with the *nptII* gene to generate bifunctional proteins that can be used as a selectable marker and as a reporter (Fig. 10.8). The greatest disadvantage is that both detection assays are destructive to the cells. The substrates are also quite expensive.

10.3.2.2. Luciferase. Whereas GUS gives transgenic cells a blue color, luciferase produces bioluminescent light (Fig. 10.9). The firefly (*Photinus pyralis*) enzyme luciferase (EC 1.13.12.7) was one of the first useful reporters for plants, though it has not been used as extensively as GUS. The enzyme catalyzes the ATP-dependent oxidative decarboxylation of luciferin as substrate. A significant advantage is the sensitive, nondestructive monitoring system that allows real-time

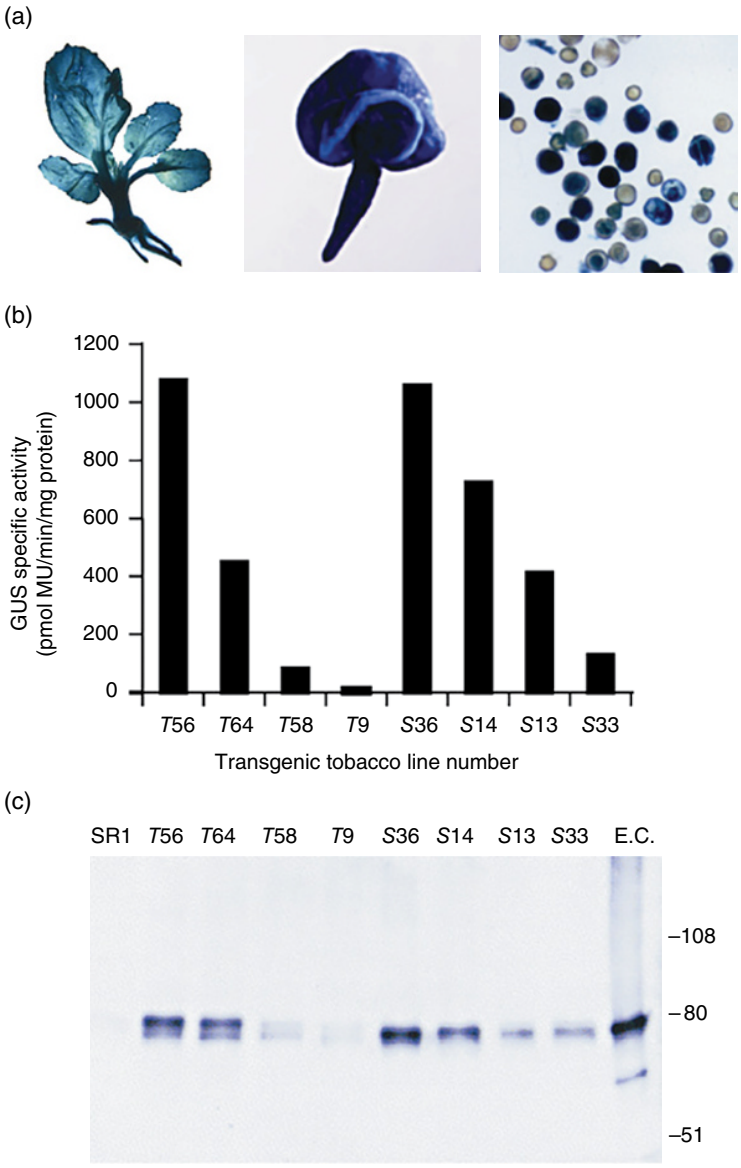


Figure 10.6. The *uidA* gene, coding for GUS, as an example of a reporter gene that has been extensively used in plants. (a) Histochemical staining for GUS activity using the substrate 4-methyl umbelliferyl glucuronide (MUG) allows detection of gene activity in specific tissues of transgenic plants. Shown in the figure are the staining of cauliflower plantlets in which constitutive expression of GUS is conferred by a strong constitutive promoter, tCUP; excised embryos from transgenic canola seeds in which seed-specific expression is conferred by the napin promoter; and transgenic canola pollen in which cell-specific expression is conferred by the pollen-specific (Bnm1) promoter. Note here that pollen cells are segregating as transformed and non-transformed cells indicated by the presence and absence of staining. (*Source*: Courtesy of Dan Brown.) (b) Measurement of GUS enzyme-specific activity using the substrate 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc). Each separate transgenic line of tobacco differs in the level of gene expression because of the variation in the influences on the inserted genes from the genetic elements and chromatin environment at the different sites of insertion. These are often called *position effects*. To compare differences among genes and elements introduced into transgenic plants, analyses must account for a large number of transgenic lines to reduce the influence of position effects. Reporter genes provide a valuable means for gathering large amounts of data. Here, a comparison of the promoter strengths of the 35S (plant lines with the *S* designation) and tCUP (plant lines with the *T* designation) constitutive promoters is inferred by comparing the activities of the reporter gene. (c) To ensure that the reporter gene reflects transcriptional activity, RNase protection assays are used to measure the relative amounts of GUS mRNA accumulating in the transgenic lines. This assay involves the formation of stable RNA duplexes with a radiolabeled antisense RNA probe followed by RNase digestion of the single-stranded RNA molecules so that the protected double-stranded RNA can be separated by gel electrophoresis and quantified. (*See insert for color representation of the figure.*)

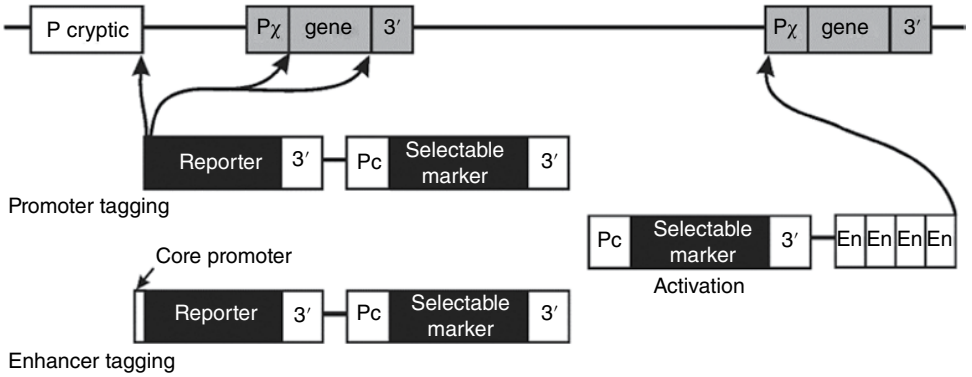


Figure 10.7. Interactions occurring between marker genes and elements in the plant genome. Various experimental strategies have been developed to probe and exploit the plant genome for functional elements and genes. This includes the use of vectors in which reporter genes are introduced into transgenic plants without key regulatory elements such as promoters and enhancers. Activation of the reporter is therefore dependent on the acquisition of the missing elements from the genome at the site of insertion. These are called *enhancer trap* or *promoter trap* experiments. The frequency of trapping such elements can be very high. The regulatory elements may be associated with expressed genes (Pχ) or may lie dormant in the genome as cryptic elements (P cryptic). An alternate strategy used to activate genes of interest (GOIs) is by introducing strong constitutive enhancer elements alone. This is often referred to as *activation tagging* of genes. Interestingly, this strategy can be combined with selection and/or screening techniques to recover genes within specific functional groups.

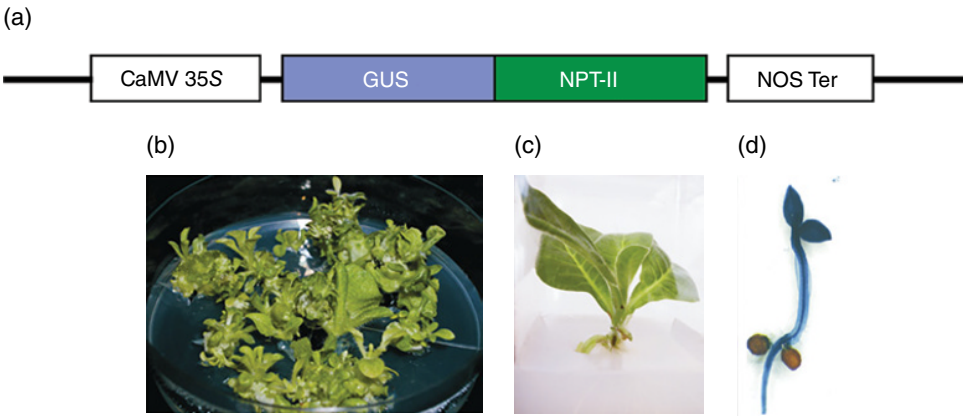


Figure 10.8. Fusion of a reporter and selectable marker gene to create a bifunctional gene: (a) GUS:NPTII fusion reporter system for plants that incorporates the *nptII* gene for kanamycin selection and the GUS reporter gene in a single module; (b) transformed tobacco shoots selected on kanamycin; (c) shoots with roots regenerated on kanamycin; and (d) a transgenic seedling after two generations showing retention of GUS gene activity indicated by the histochemical staining with the GUS substrate X-Gluc. (*Source*: From courtesy of Raju Datla.) (*See insert for color representation of the figure.*)

analysis. Furthermore, the half-life of the luciferase protein in plant cells is lower than that for GUS and may reflect transcriptional activity more accurately. It is often used as an internal control in experiments that require the use of more than one reporter system.

10.3.2.3. Fluorescent Proteins. FPs have become the most important reporter gene system for plants. They require no external substrate for detection, and there have been no reports of detrimental effects on the fitness of plants that express them. The novelty of FPs is that they combine great sensitivity at the subcellular level using bioimaging technologies made available through

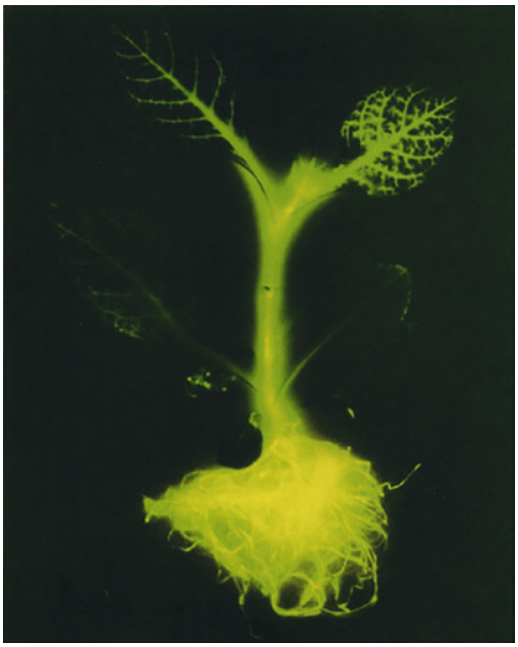


Figure 10.9. Luminescence detected in transgenic tobacco transformed with the firefly luciferase gene driven by the 35S promoter and watered with a solution of luciferin, the luciferase substrate. (*Source*: From Ow et al. (1986). Reproduced with permission of AAAS.) (*See insert for color representation of the figure.*)

confocal laser scanning microscopy with real-time detection in living cells (Fig. 10.10). A wide variety of FPs are being developed to extend the range and complexity of processes that can be simultaneously monitored in living cells (Stewart 2006). Among them, the Pacific jellyfish (*Aequorea victoria*) green FP (GFP) is the most widely used FP in plants, which glows green under UV light. New FPs are being discovered, and we can expect a rainbow of colors in the near future (Stewart 2006). Especially useful are FPs that emit in the orange and red spectra such as an orange FP (OFP) *pporRFP* from the hard coral *Porites porites* and a synthetic OFP mOrange from *Discosoma* sp. (Mann et al. 2012; Fig. 10.11). Their usefulness is because natural auto-fluorescence is less in the orange and red spectra in most plants. These FPs can be fused to various plant proteins and used as a tag to monitor their trafficking and interactions (Fig. 10.10). In field studies, FPs also permit the rapid and easy detection of transgenic plants or plant parts such as pollen (Fig. 10.12). In tissue culture, FPs have been used in combination with selectable marker genes to identify and enrich the content of transformed material to improve the recovery of transgenic plants from species where the current transformation and selection systems are inefficient.

10.4. MARKER-FREE STRATEGIES

Because selectable marker genes are often only utilized in the generation of transgenic plants and not after transgenic plants have been developed, there is generally no need to maintain them in the transgenic plants. Their removal may provide some advantages if the plant is to be used for another round of transformation because the same selectable marker gene, if effective, can be used repeatedly. If the safety of the selectable marker to health or the environment is a concern, it may be useful to have a method to remove it from the plant before commercialization. Furthermore, it would be essential to remove the selectable marker gene if it alters plant growth and differentiation. In rare cases, the transformation frequency may be high enough to recover transgenic plants

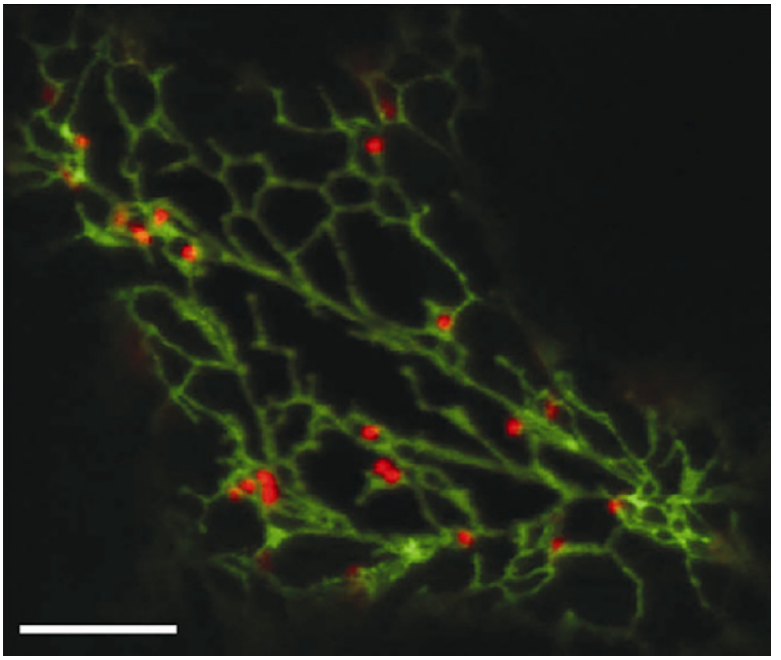


Figure 10.10. Confocal laser scanning microscopy of leaf mesophyll cells transiently expressing peptides fused to green fluorescent protein or GFP (green image) and yellow fluorescent protein (red image). GFP is fused to the HDEL tetrapeptide (spGFP-HDEL) to achieve ER retention and thus reveals the cortical ER network in leaf cells. The proximity of the Golgi to the ER network is revealed by the yellow FP fused to a Golgi glycosylation enzyme (ST-YFP). (Bar = 10 μm.) (*Source*: From Brandizzi et al. 2004.) (*See insert for color representation of the figure.*)

through screening techniques without selection. With certain traits, such as herbicide resistance, the GOI may be directly selectable without the need for a separate selectable marker gene (Darbani et al. 2007).

The easiest method for generating marker-free plants is the co-transformation of the GOI with a marker gene followed by segregation of the unlinked genes into separate lines (Fig. 10.13). Although effective, this requires the production of many transgenic lines initially. Furthermore, the technology is restricted to transgenic plants that are propagated through seeds. This would exclude vegetatively propagated species, such as trees. Co-transformation can be achieved in many ways. For instance, the two genes can be introduced in two separate plasmids. If *Agrobacterium*-mediated transformation is used, this can be achieved by infecting tissues with separate plasmids using separate *Agrobacterium* strains or by separate plasmids in one strain. Co-transformation could also be achieved using a single strain carrying a single plasmid with two separate T-DNA regions. The frequencies of co-transformation may be very high (>50%, depending on the situation), allowing the selection of transgenic material carrying both the selectable marker and the gene of interest. The segregation of transgenic lines carrying the GOI from lines carrying the selectable marker gene may occur at frequencies sufficiently high to be practical in species that are efficiently transformed.

Only a few studies have emerged in which transposons have been used to translocate genes within the plant genome to break the linkage between selectable marker genes and GOIs (Fig. 10.13). An interesting strategy has been developed using the selectable marker gene *ipt* in combination with the Ac transposase element to remove the *ipt* gene. In this multiautotransformation (MAT) process, the *ipt* gene first acts positively to generate a proliferation of morphologically abnormal shoots with a "shooty" phenotype. They cannot regenerate because of the overproduction of cytokinin; however, normal shoots emerge at low frequency several weeks later following transposition of the *ipt* gene

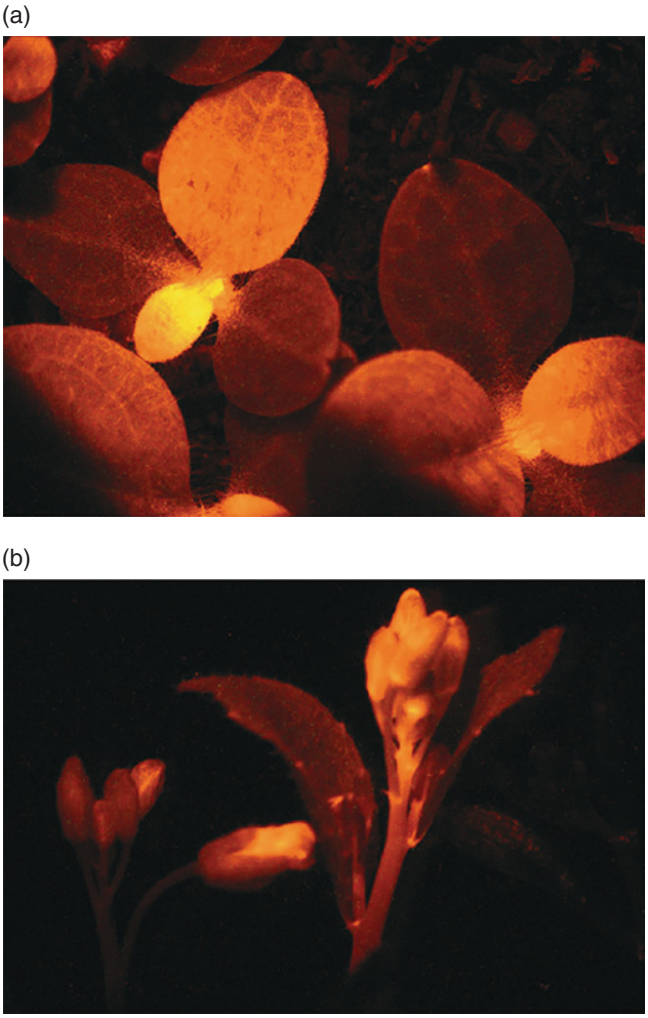


Figure 10.11. Orange fluorescent proteins whose genes were cloned from corals and expressed in tobacco (a) and *Arabidopsis* (b) plants. (*See insert for color representation of the figure.*)

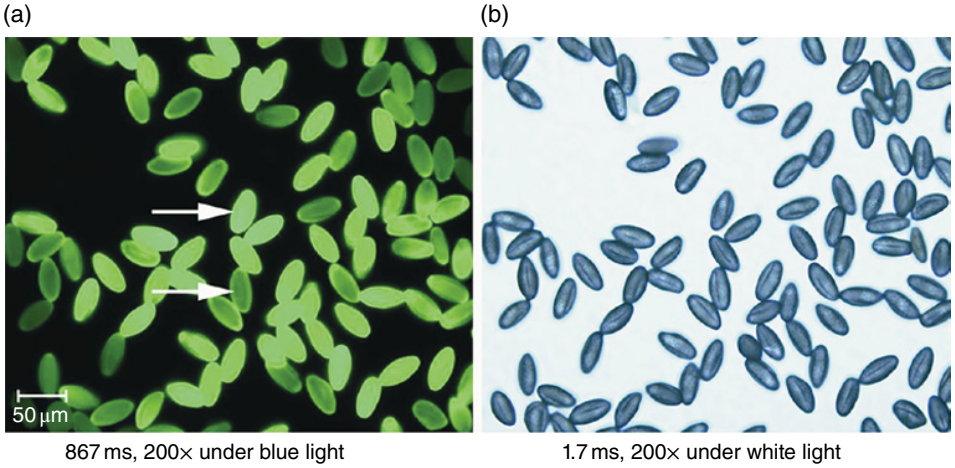


Figure 10.12. The green fluorescent protein (GFP) has been useful for marking whole plants using a 35S-GFP construct and plant parts such as pollen using GFP under the control of a pollen-specific promoter (Lat59) from tomato: (a) 867 ms, 200× under blue light and (b) 1.7 ms, 200× under white light. The arrows in (a) show GFP fluorescence of pollen cells. (*Source*: Courtesy of Moon & Stewart.) (*See insert for color representation of the figure.*)

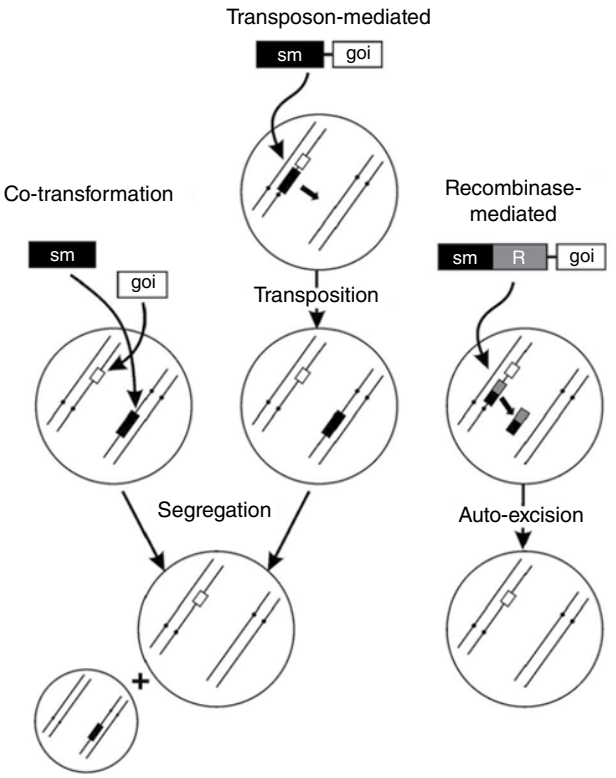


Figure 10.13. Processes for generating marker-free transgenic plants. Co-transformation is a practical process for generating marker-free transgenic plants. It depends on the integration of the selectable marker gene (sm) and the gene of interest (GOI) at separate chromosomal sites that can segregate away from each other in the next sexual generation. This can also be achieved by the introduction of both the selectable marker and the GOI on the same vector and therefore insertion at the same site followed by the subsequent transposition of the marker gene to a separate locus that can segregate away from the GOI. A more recent advance is the use of an excision recombinase (R) under the control of an inducible promoter which autoexcises itself along with the selectable marker gene. This process does not require a segregation step and has the potential for broader applications.

to a distant locus that can segregate away in somatic cells or it may be directly lost if not reinserted into the genome.

More recently, the use of site-specific recombinases has emerged as a versatile strategy for the selective removal of marker genes from an insertion site. The recombinases and their target sites include Cre/*lox* from bacteriophage P1, FLP/*FRT* from yeast *Saccharomyces cerevisiae*, R/*RS* from *Zygosaccharomyces rouxii*, CinH/*RS2* from *Acetinetobacter*, and Par/*MRS* from a broad-host-range plasmid RP4. The target sequences are placed around the genes targeted for excision followed by the introduction of the recombinase in a second round of transformation. Again, this approach suffers from the problem that it is restricted to seed-propagated plants in order to segregate the recombinase gene from the gene of interest. This has been partially overcome by the introduction of the recombinase via transient expression. Although excision occurs at a lower frequency, the recombinase gene is not integrated into the genome. Another promising approach incorporated the GOI along with the selectable marker genes and recombinase gene on one vector (Fig. 10.13). This strategy overcomes many of the earlier limitations by using an inducible promoter to express the recombinase, resulting in the autoexcision of the recombinase and the selectable marker genes simultaneously. This approach eliminates the need for successive rounds of transformation or crossing and minimizes the

period of exposure of plants to the action of recombinases. Prolonged exposure to recombinases is a concern as unpredictable deletions in the genome may occur due to the action of recombinases on cryptic target sites. Although the extent of such deletions is uncertain in the nuclear genome, examples have been reported in the plastid genome with Cre. With time, the excision systems and gene regulatory systems will improve, and the technology is likely to be refined to practical levels. The strategies discussed in the text as well as other methods for marker gene removal in plants have been reviewed by Sang et al. (2013) and Yau and Stewart (2013).

10.5. CONCLUSIONS

To obtain a better understanding of gene functions in both basic and applied studies, transgenes should be precisely regulated. Promoters play a key role in the regulatory of transgene expression. Transgenes can be driven by constitutive, tissue-specific, or inducible promoters. Promoter activities are determined by the interactions between promoter motifs and corresponding transcription factors. Considering the huge number of plant genes, it would not be surprising that plant promoters and motifs remain to be discovered and functionally characterized. Advances in *de novo* motif discovery will lead to increased discovery of novel promoter motifs. Engineering of promoter motifs for the design and construction of synthetic promoters is expected to play a critical role in promoter engineering in the near future.

The production of transgenic plants is very difficult without the use of marker genes. As we have seen, effective marker genes have virtually no effect on plant phenotype except their intended effect. Nonetheless, marker genes have been somewhat controversial, especially antibiotic resistance genes, because of the concern about horizontal gene transfer (HGT). HGT is the movement of DNA from one species to an unrelated species—in this case, from transgenic plants to bacteria. In the event that an antibiotic resistance gene were to be horizontally transferred to bacteria, some people worry that new antibiotic resistance problems could be created that could harm human or ecosystem health. Even though HGT has not been demonstrated from transgenic plants to bacteria in a realistic experimental system, it has affected the politics of regulation and the perception of transgenic plants, which will be covered in later chapters.

As we saw in Chapter 9, when constructing transformation vectors, a number of factors must be considered in addition to the promoter selected to drive the expression of marker genes. For example, the orientation of the promoter within the transferred DNA is also extremely important. The 35S promoter may interact with neighboring promoters in the vector and plant sequences at the insertion site. It is known to radically alter the specificity of tissue-specific promoters. Field studies have also shown that interactions could occur with infecting viral sequences resulting in transgene silencing. For many years, there have been observations of gene silencing mediated through sequence similarity between promoters introduced on transformation vectors and resident genes. For example, inverted repeats within the DNA sequence must be avoided (Fig. 10.3). Furthermore, the promoters associated with the selectable markers are located at maximal distances from promoters associated with the GOI or the borders with the plant DNA (Fig. 10.14). The genome has an abundance of promoters that could interact with incoming DNA on insertion. Indeed, this knowledge has been used for promoter discovery research through the use of promoter or enhancer trap strategies (Fig. 10.7). Interestingly, a new strong constitutive promoter, tCUP (Fig. 10.6), was discovered in this way and was found to be very useful for driving the expression of selectable marker genes because it did not interact with other promoters as extensively as the 35S promoter (Fig. 10.14). Reporter genes have been used extensively to study the specificity and level of plant promoter activity; therefore, the promoters combined with reporter genes have been much more diverse. As discussed earlier, the orientation and type of promoter fused to the reporter gene must be carefully balanced, with the promoter fused to the selectable marker genes (Fig. 10.14).

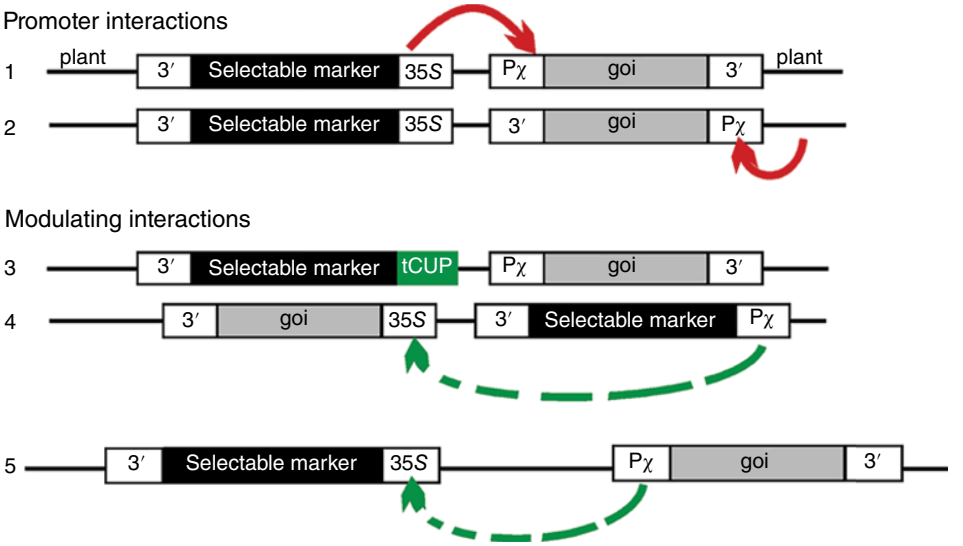


Figure 10.14. Construction of plant transformation vectors to avoid interactions among promoters used to drive selectable marker genes and genes of interest (GOIs). Constitutive promoters, such as the 35S promoter, are frequently used to drive expression of the selectable marker genes. However, the 35S promoter will interact with other promoters (solid arrows) within the transformation vector, particularly if they are situated near each other (configuration 1). This can lead to aberrant or unpredictable expression of the GOI. Similar interactions may occur with elements within the plant DNA that become positioned close to promoters within the transferred DNA (configuration 2). These interactions can be minimized by the design of the vector. The simplest approach is to use a constitutive promoter that does not tend to interact with other promoters, for example, the tCUP promoter (shown in configuration 3). The genes within the transferred DNA can also be positioned so that their promoters are spaced as far away from each other as possible through their orientation relative to each other (configuration 4) or by the insertion spacer DNA between them (configuration 5). These manipulations will reduce the extent of the interactions (indicated by the broken arrows).

LIFE BOX 10.1. FREDY ALTPETER

Fredy Altpeter, Research Foundation Professor, Department of Agronomy, University of Florida—IFAS.



Fredy Altpeter exploring caves in the British Virgin Islands. Courtesy of Fredy Altpeter.

I grew up in Riegelsberg, a small village in the Southwest of Germany, and was the first of my family to attend college. Actually, I was not really sure if I should go to college, considering my minimalistic approach in high school and the resulting mediocre grades. I, like all Germans, had to complete mandatory military service of 18 months at the time when I graduated from high school. While I enjoyed the camaraderie, the daily routine in the Falkenstein barracks sparked a desire for higher studies. I remember spending a lot of time reading and thinking about different career paths and was enthused about studying agricultural science and its global importance.

Applied aspects of science were one of the few topics that got my attention in school and

seemed to be the focus of the coursework for agricultural sciences. My initial career goal was to become a consultant for farmers. Since I did not grow up on a farm, I decided to complete a vocational degree as farm manager to "speak the same language as farmers" and gain relevant practical experience before college. Working for 2 years on different farms paid off later in many ways including the ability to troubleshoot technical problems as well as a keen interest in developing and applying strategies for crop improvement. I went on to attend the University of Hohenheim in Germany, and was excited about the variety of applied courses ranging from plant science over farm machinery technology to economics. I was most fascinated by plant breeding and plant pathology and therefore majored in plant science. I was impressed by a plant biotechnology lecture series from Gerd Weber in which he challenged us with the technical complexity of protoplast-mediated gene transfer and its molecular characterization.

My first research project focused on the development of a protocol for production and purification of mycotoxins with the goal to understand their role in pathogenicity. This field of research was even new to my advisor and grass breeder Ulrich Posselt. When initial attempts failed, I applied for travel funding to support a short internship in David Miller's mycotoxin research group at the NRC in Ottawa, Canada. This experience not only put my research project on the path of success, but it was also a great social and travel experience that left me with a desire to return to North America later in my career. I decided to stay at the Plant Breeding and Biotechnology Center of the University of Hohenheim to work on my PhD research project with Ulrich Posselt and Gerd Weber. I developed cell culture regeneration protocols for perennial ryegrass and use them for *in vitro* selection of mycotoxin-resistant lines.

This relatively new area of plant biotechnology was fast-moving and I was impressed with Indra Vasil's publication describing the first transgenic wheat plants. I was very happy that Indra accepted me as a postdoctoral research associate in his wheat biotechnology program at the University of Florida. After developing a routine wheat transformation protocol, we generated wheat plants with improved bread-making quality and insect resistance. Gainesville, the hometown of the University of Florida, was also a great place to live, and my wife Angelika and I had fun exploring many nearby tourist attractions and beaches. But we were not yet ready to settle in the United States and accepted group leader positions at the IPK Gatersleben, Germany. The IPK Gatersleben provided an excellent research infrastructure, which made it easier to compete for external funding and allowed me to establish a cereal and grass transformation group to address a number of biological questions ranging from biotic stress to bread-making quality. I had the good fortune of collaborating with many eminent scientists at the IPK and being able to train great technicians, graduate students, postdocs, and visiting scientists.

However, it would not take long until my wife and I were missing our friends, the sunny skies, beaches, and people in Florida. The call from Jerry Bennett, chair of the Agronomy Department at the University of Florida, offering me a faculty position was an answered prayer and I was ecstatic to accept it. Building on my previous experience, I developed a broad research program integrating translational genomics, genetic transformation, plant breeding, and newly emerging technologies, such as genome editing, to improve cereals, grasses, and biofuel feedstocks. Biological questions addressed under controlled environment and field conditions include abiotic and biotic stress, photosynthetic efficiency to enhance biomass yield, turf quality, risk assessment and management, cell wall composition for improved conversion to biofuel and forage digestibility, and metabolic engineering for production of oil in vegetative tissues.

Being actively involved in professional societies, such as the Society for In Vitro Biology (SIVB) and the Crop Science Society of America (CSSA) facilitated networking, set up collaborations, identification of sponsors, and recruitment of talented graduate students and postdocs. The basic and applied aspect of my research attracted not only federal funds from NSF, DOE, and USDA but also industry

support from the Scotts Company, Vialactia and Syngenta. The interactions with excellent collaborators from academia and industry like Steve Long, Roger Wise, Don Ort, John Shanklin, Maria Gallo, Jim Preston, Ken Quesenberry, Kevin Kenworthy, Wilfred Vermerris, John Erickson, Ann Blount, Yan Zhang, Shujie Dong, Kasi Azhakanandam, Aron Silverstone, Ian Jepson, Sathish Puthigae, and Bob Harriman were essential in refining strategies to push boundaries and maintaining the high level of excitement that makes this a rewarding experience for me and the many outstanding students, postdocs, and visiting scientist in my program. Currently, we are working at the forefront of synthetic biology including precision genome editing, metabolic engineering, and regulatory switches.

Based on my personal experience, my advice to young scientist is to complete multiple internship experiences at an early stage of your career. This will help you to carefully choose the research field that you are most passionate about. Your excitement will give you the ideas, energy, and persistence which will lead you to success. It is also important to balance your life and recharge outside of the lab. I love Christ and enjoy going to church since college as well as outdoor activities with family including hiking, cross country skiing, boating, swimming, and beach walking.

LIFE BOX 10.2. TANIYA DHILLON

Taniya Dhillon, Postdoctoral Associate, University of Tennessee.



Taniya Dhillon. Courtesy of Taniya Dhillon.

For as long as I can remember, I have lived in an environment influenced by agriculture. My father's family practiced farming in Punjab, the major agricultural state of India. Both my parents were professors in soil science. We lived on the university campus, where I spent the first 15 years of my life growing up, watching fields of wheat and rice surrounding me in alternating seasons. I believe this ambience and the close connection with agriculture and science played a huge role in shaping my scientific curiosity.

I followed my passion and got a bachelor's degree in agriculture with honors in crop science in India. While studying the vast variety of subjects encompassed within agriculture, I realized that plant breeding, genetics, and biotechnology were the most interesting to me. I loved wheat maybe because I consumed it every day or probably because its history is rather impressive, going from diploid, to tetraploid, and finally to present-day hexaploid. Seeing the variability within wheat plants was very exciting to me. The question I had was "why"—why did some plants have awns while others didn't, why their height and color were so variable, why some plants' ears bore more seeds than other plants, when they all belonged to the same wheat species. The fact that this variation was genetically controlled was very intriguing. That there was a gene or genes, whose nucleotide sequence determined a plant's appearance was very interesting and still is, it's just that over time I've learned that a plant's or any organism's phenotype is a manifestation of the interplay

between several factors than simply a gene's nucleotide sequence.

I decided to pursue a career in plant sciences and came to the United States after I was admitted to the master's program at The Ohio State University. Little did I know at the time that I would be spending the next 8 years at the OARDC campus in Wooster, working toward a masters and following it up with a PhD. My masters research specialization was in biotechnology in Dr. John Finer's lab. Our objective was to find ways to stabilize transgene expression in plants. We conducted this research using lima bean cotyledons and soybean embryogenic tissue. I enjoyed cloning DNA sequences into plasmids, introducing the plasmid constructs into plants, screening the resulting transgenics, and then monitoring the change in their phenotype for the better, although sometimes for the worse. I realized during this time that I loved doing basic research, and my goal was to apply this basic research in a way that would benefit farmers and mankind in general.

While working on my masters, I learned of the research being carried out in a neighboring lab led by Dr. Eric Stockinger, where they were working on understanding the molecular genetic mechanisms that enable a plant to survive freezing temperatures in the winter. I got the opportunity to work with them to fulfill a graduation requirement, and enjoyed this work so much that I made this lab my home for a PhD. Our focus was on the Triticeae crops, wheat, barley, and rye, which provided us the potential to help farmers growing these crops in places such as Ukraine and Great Plains of the United States, where the winter is usually very severe. Through this research, I was able to expand my skillset into the areas of molecular biology and plant physiology. We discovered that one of the mechanisms underlying freezing tolerance was copy number variation; simply put, while all plants possessed the genes that impart freezing tolerance, a plant with greater copy numbers of these genes was more freezing-tolerant than the one with fewer copies. Artificially increasing the copy numbers of these genes, however, had undesirable effects, one of which was delayed flowering. We explored this further and demonstrated the connection between freezing tolerance and flowering time, and the genes and gene networks involved.

Following PhD, I secured a postdoctoral position in the lab of Dr. Andrew Flavell at the University of Dundee, James Hutton Institute (JHI), Scotland, UK. JHI is one of the world's leading institutes in barley research. Here, some of the latest techniques such as ChIP-Seq and exome capture are used to study barley. ChIP-Seq examines the epigenome that is, the genome-wide distribution of histone modifications, to help understand the chromosomal features that define the condensed heterochromatin from the relaxed euchromatin, which has further implications in genetic recombination and crop improvement. The other approach, exome capture, allows mining the genetic variability in the protein coding regions or exons of the genome. We have used this technique to identify the genetic variation in wild barley that can be introgressed into cultivated barley to improve its adaptability to the changing global climate. This research work has helped me develop computational and bioinformatics skills, which is highly valued in today's scientific research.

Currently, I'm in Dr. Neal Stewart's lab with a postdoctoral position at the University of Tennessee (UT). I met Dr. Stewart at a conference when I was at Ohio State. The opportunity to return to the United States was to continue developing my interests in molecular biology and biotechnology. My research project at UT is to develop high-throughput single-cell suspension culture system in switchgrass and maize for genome editing. This work is again very exciting to me as it allows me to expand my research experience to new crops. It also has tremendous application potential; the ultimate goal of this research is to improve the efficiency of biofuel production from these crops, which can significantly boost the biofuel industry and also help save the environment by reducing fossil fuel consumption.

Throughout my academic career, I've enjoyed working with great people in varied fields of plant science in different labs and across different countries. My goal is to secure a permanent position at a university and continue working toward improving plants for humankind.

REFERENCES

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997): Role of *Arabidopsis* MYC and MYB homologs in drought-and abscisic acid-regulated gene expression. *Plant Cell* **9**:1859–1868.

Aoyama T, Chua N-H (1997): A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J* **11**:605–612.

Bhullar S, Chakravarthy S, Advani S, Datta S, Pental D, Burma PK (2003): Strategies for development of functionally equivalent promoters with minimum sequence homology for transgene expression in plants: *cis*-elements in a novel DNA context versus domain swapping. *Plant Physiol* **132**:988–998.

Brandizzi F, Irons SL, Johansen J, Kotzer A, Neumann U (2004): GFP is the way to glow: bioimaging of the plant endomembrane system. *J Microsc* **214**:138–158.

Bruce W, Folkerts O, Garnaat C, Crasta O, Roth B, Bowen B (2000): Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P. *Plant Cell* **12**:65–79.

Caddick MX, Greenland AJ, Jepson I, Krause K-P, Qu N, Riddell KV, Salter MG, Schuch W, Sonnewald U, Tomsett AB (1998): An ethanol inducible gene switch for plants used to manipulate carbon metabolism. *Nat Biotechnol* **16**:177–180.

Cazzonelli CI, Velten J (2008): *In vivo* characterization of plant promoter element interaction using synthetic promoters. *Transgenic Res* **17**:437–457.

Chaturvedi CP, Sawant SV, Kiran K, Mehrotra R, Lodhi N, Ansari SA, Tuli R (2006): Analysis of polarity in the expression from a multifactorial bidirectional promoter designed for high-level expression of transgenes in plants. *J Biotechnol* **123**:1–12.

Christensen HEM, Ramachandran S, Tan C-T, Surana U, Dong C-H, Chua N-H (1996): *Arabidopsis* profilins are functionally similar to yeast profilins: identification of a vascular bundle-specific profiling and a pollen-specific profiling. *Plant J* **10**:269–279.

Comai L, Moran P, Maslyar D (1990) Novel and useful properties of a chimeric plant promoter combining CaMV 35S and MAS elements. *Plant Mol Biol* **15**:373–381.

Cordero MJ, Raventos D, Segundo BS (1994): Expression of a maize proteinase inhibitor gene is induced in response to wounding and fungal infection: systemic wound-response of a monocot gene. *Plant J* **6**:141–150.

Corrado G, Karali M (2009): Inducible gene expression systems and plant biotechnology. *Biotechnol Adv* **27**:733–743.

Darbani B, Eimanifar A, Stewart CN Jr, Camargo W (2007): Methods to produce marker-free transgenic plants. *Biotechnol J* **2**:83–90.

De Almeida ERP, Gossele V, Muller CG, Dock J, Reynaerts A, Botterman J, Krebbers E, Timko MP (1989): Transgenic expression of two marker genes under the control of an Arabidopsis rbcS promoter: sequences encoding the Rubisco transit peptide increase expression levels. *Mol Gen Genet* **218**:78–86.

Ebinuma H, Sugita K, Matsunaga E, Endo S, Yamada K, Komamine A (2001): Systems for the removal of a selectable marker and their combination with a positive marker. *Plant Cell Rep* **20**:383–392.

Flavell RB, Dart E, Fuchs RL, Fraley RT (1992): Selectable marker genes: safe for plants. *Biotechnology* **10**:141–144.

Gatz C, Frohberg C, Wendenburg R (1992): Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. *Plant J* **2**:397–404.

Guerineau F, Benjdia M, Zhou DX (2003): A jasmonate-responsive element within the *Arabidopsis thaliana* vsp1 promoter. *J Exp Bot* **54**:1153–1162.

Guiltinan MJ, Marcotte WR, Quatrano RS (1990): A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* **250**:267–71.

Guo H, Chen X, Zhang H, Fang R, Yuan Z, Zhang Z-S, Tian Y (2004): Characterization and activity enhancement of the phloem-specific pumpkin PP2 gene promoter. *Transgenic Res* **13**:559–566.

Hamilton DA, Schwarz YH, Mascarenhas JP (1998): A monocot pollen-specific promoter contains separable pollen-specific and quantitative elements. *Plant Mol Biol* **38**: 663–669.

Hattori T, Terada T, Hamasuna S (1995): Regulation of the *Osem* gene by abscisic acid and the transcriptional activator VP1: analysis of *cis*-acting promoter elements required for regulation by abscisic acid and VP1. *Plant J* **7**:913–925.

Jeong JS, Kim YS, Baek KH, Jung H, Ha S-H, Choi YD, Kim M, Reuzeau C, Kim J-K (2000): Root-specific expression of *OsNAC10* improves drought tolerance and grain yield in rice under field drought conditions. *Plant Physiol* **153**:185–197.

Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB (1990): Different temporal and spatial gene expression patterns occur during anther development. *Plant Cell* **2**:1201–1224.

Ku MSB, Agarie S, Nomura M, Fukayama H, Tsuchida H, Ono K, Hirose S, Toki S, Miyao M, Matsuoka M (1999): High-level expression of maize phosphoenolpyruvate carboxylase in transgenic rice plants. *Nat Biotechnol* **17**:76–80.

Lam E, Chua N-H (1991): Tetramer of a 21-base pair synthetic element confers seed expression and transcriptional enhancement in response to water stress and abscisic acid. *J Biol Chem* **266**:17131–17135.

Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E (1998): Functional analysis of regulatory sequences controlling PR-1 gene expression in *Arabidopsis*. *Plant J* **16**:223–233.

Li ZT, Jayasankar S, Gray DJ (2004) Bi-directional duplex promoters with duplicated enhancers significantly increase transgene expression in grape and tobacco. *Transgenic Res* **13**:143–154.

Lois R, Dietrich A, Hahlbrock K, Schulz W (1989): A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive *cis*-acting elements. *EMBO J* **8**:1641–1648.

Luschnig C, Gaxiola RA, Grisafi P, Fink GR (1998): EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in *Arabidopsis thaliana*. *Genes Dev* **12**:2175–2187.

Mann DGJ, Abercrombie LL, Rudis MR, Millwood RJ, Dunlap JR, Stewart CN Jr (2012): Very bright orange fluorescent plants: endoplasmic reticulum targeting of orange fluorescent proteins as visual reporters in transgenic plants. *BMC Biotechnol* **12**:17.

Mariani C, De Beuckeleer M, Truettner J, Leemans J, Goldberg RB (1990): Induction of male sterility in plants by a chimaeric ribonuclease gene. *Nature* **347**:737–741.

Miki B, McHugh S (2004): Selectable marker genes in transgenic plants: applications, alternatives and biosafety. *J Biotechnol* **107**:193–232.

Mitra A, Choi HK, An G (1989): Structural and functional analyses of *Arabidopsis thaliana* chlorophyll a/b-binding protein (*cab*) promoters. *Plant Mol Biol* **12**:169–179.

Muller B, Sheen J (2008): Cytokinin and auxin interplay in root stem-cell specification during early embryogenesis. *Nature* **453**:1094–1097.

Ni M, Cui D, Einstein J, Narasimhulu S, Vergara CE, Gelvin SB (1995): Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. *Plant J* **7**:661–676.

Orozco BM, Ogren WL (1993): Localization of light-inducible and tissue-specific regions of the spinach ribulose bisphosphate carboxylase/oxygenase (rubisco) activase promoter in transgenic tobacco plants. *Plant Mol Biol* **23**:1129–1138.

Ow D, Wood KV, DeLuca M, DeWey JR, Helinski DR, Howell SH (1986): Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* **234**:856–859.

Puente P, Wei N, Deng XW (1996): Combinational interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in *Arabidopsis*. *EMBO J* **15**:3732–3743.

Rask L, Ellerstrom M, Ezcurra I, Stalberg K, Wycliffe P (1998): Seed-specific regulation of the Napin promoter in *Brassica napus*. *J Plant Physiol* **152**:595–599.

Rushton PJ, Reinstedler A, Lipka V, Lippok B, Somssich IE (2002): Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen-and wound-induced signaling. *Plant Cell* **14**:749–762.

Sang Y, Millwood RJ, Stewart CN Jr. (2013): Gene use restriction technologies for transgenic plant bioconfinement. *Plant Biotechnol J* **11**:649–658.

Sawant S, Singh PK, Madanala R, Tuli R (2001): Designing of an artificial expression cassette for the high-level expression of transgenes in plants. *Theor Appl Genet* **102**:635–644.

Shah J, Klessig DF (1996): Identification of a salicylic acid-responsive element in the promoter of the tobacco pathogenesis-related-1,3-glucanase gene, *PR-2d*. *Plant J* **10**:1089–1101.

Shen Q, Uknes SJ, Ho T-HD (1993): Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene. *J Biol Chem* **268**:23652–23660.

Simpson J, Timko MP, Cashmore AR, Schell J, Montagu MV, Herrera-Estrella L (1985): Light-inducible and tissue-specific expression of a chimaeric gene under control of the 5′-flanking sequence of a pea chlorophyll a/b-binding protein gene. *EMBO J* **4**:2723–2729.

Stewart CN Jr (2006): Go with the glow: fluorescent proteins to light transgenic organisms. *Trends Biotechnol* **24**:155–162.

Stockinger EJ, Gilmour SJ, Thomashow MF (1997): *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc NatL Acad Sci USA* **94**:1035–1040.

Thornburg RW, An G, Cleveland TE, Johnson R, Ryan CA (1987): Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase gene fusion in transgenic tobacco plants. *Proc Natl Acad Sci USA* **84**:744–748.

Tran LSP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2004): Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-response *cis*-element in the *early responsive to dehydration stress 1* promoter. *Plant Cell* **16**:2481–2498.

Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S (1991): Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Dev* **5**:496–507.

Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ (1997): Aux/IAA protein repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**:1963–1971.

Unger E, Cigan AM, Trimnell M, Xu R-J, Kendall T, Roth B, Albertsen M (2002): A chimeric ecdysone receptor facilitates methoxyfenozide-dependent restoration of male fertility in *ms45* maize. *Transgenic Res* **11**:455–465.

Weisshaar B, Armstrong GA, Block A, da Costa e Silva O, Hahlbrock K (1991): Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promoter element with functional relevance in light responsiveness. *EMBO J* **10**:1777–1786.

Wu C-Y, Washida H, Onodera Y, Harada K, Takaiwa F (2000): Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene. *Plant J* **23**:415–421.

Xie M, He Y, Gan S (2001): Bidirectionalization of polar promoters in plants. *Nat Biotechnol* **19**:677–679.

Yamaguchi-Shinozaki K, Shinozaki K (1994): A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**:251–264.

Yau YY, Stewart CN Jr (2013): Less is more: strategies to remove marker genes from transgenic plants. *BMC Biotechnol* **13**:36.

Yu D, Chen C, Chen Z (2001): Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR-1* gene expression. *Plant Cell* **13**:1527–1539.

Zarka DG, Vogel JT, Cook D, Thomashow MF (2003): Cold induction of *Arabidopsis CBF* genes involves multiple ICE (inducer of *CBF* expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiol* **133**:910–918.

Zuo J, Niu Q-W, Chua N-H (2000): An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* **24**:265–273.

Zuo J, Niu Q-W, Moller SG, Chua N-H (2001): Chemical-regulated, site-specific DNA excision in transgenic plants. *Nat Biotechnol* **19**:157–161.

CHAPTER 11

Transgenic Plant Production

JOHN J. FINER

Department of Horticulture and Crop Science, OARDC/The Ohio State University, Wooster, Ohio

11.0. CHAPTER SUMMARY AND OBJECTIVES

11.0.1. Summary

Although DNA can be introduced into plant cells using a variety of creative approaches, the two most commonly used methods employ *Agrobacterium tumefaciens* and particle bombardment. *Agrobacterium* is the method of choice and relies on a natural genetic engineer: the causal agent of crown gall disease in plants, which is one of the most intriguing stories in plant pathology. Particle bombardment relies on accelerating DNA-coated microscopic particles into plant cells to deliver DNA to the genome. There are other less often used methods and plenty of potential technological development opportunities to improve transformation efficiency.

11.0.2. Discussion Questions

1. What is a transgene? A transgenic plant?
2. What part or parts of the plant cell provide the most resistance to DNA introduction?
3. In the case of a successful DNA introduction, where in the target cell does the foreign DNA end up?
4. What are some differences between physical and biological methods for DNA introduction into plant cells?
5. What are some ways that the biological method for DNA introduction (*Agrobacterium*) has been improved over the years?
6. How is gene introduction performed with the model plant, *Arabidopsis*? Is this technique widely applied to other plants?
7. What are the size and composition of the particles, which are used for the particle bombardment method?
8. How do the DNA integration patterns differ in plant cells, transformed via *Agrobacterium* and particle bombardment?
9. Can you think of additional methods for DNA introduction into plant cells?
10. Why do scientists sometimes "rush to publish" their research results, and why can this be problematic?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

11.1. OVERVIEW OF PLANT TRANSFORMATION

11.1.1. Introduction

Transgenic plants can be simply defined as plants that contain additional or modified genes, which were introduced using specific physical or biological methods. The introduced DNAs or transgenes are typically very well-defined and are precisely manipulated in the laboratory prior to delivery into the target plant cells. The methods for DNA introduction into plants cells are quite varied and are largely dependent on the plant selected for study and the background of the scientist performing the work. Over the years, tremendous efforts have been placed in development of gene introduction or "transformation" technology and, for many if not most plants, the procedures have become routine. For even the most difficult-to-transform plants, successful DNA introduction can often be consistent but inefficient. Fortunately, the efficiency of transgenic plant production is still being improved and new methods for DNA delivery and modification are still being investigated.

Although numerous methods have been developed for production of transgenic plants, *Agrobacterium* and particle bombardment are the two main methods used by most transformation laboratories. *Agrobacterium* has been called a "natural genetic engineer" and relies on this biological vector for transgene introduction. Particle bombardment is a physical method for DNA delivery and utilizes DNA-coated microscopic particles, which are accelerated toward a suitable target tissue. Although almost all laboratories are now using *Agrobacterium* for production of transgenic plants, most of the early transgenic crops were produced via particle bombardment. Other procedures for DNA delivery do exist and each has benefits and drawbacks. In order to better understand the challenges of producing transgenic plants and the overall process, one must first try to visualize DNA delivery to a single target plant cell and have a basic understanding of how to recover a whole genetically engineered plant from that single targeted cell.

11.1.2. Basic Components for Successful Gene Transfer to Plant Cells

11.1.2.1. Visualizing the General Transformation Process. Prior to the first successful production of transgenic plants in the mid-1980s (Horsch et al. 1985), efforts to improve plants relied exclusively on classical plant breeding through sexual hybridization and evaluation of spontaneous or induced mutations. Although plant breeding remains the foundation of plant improvement, a typical sexual cross results in the mixing of tens of thousands of genes and requires sorting through progeny to find the individuals that contain the traits or markers of interest. Through transgenic plant production, genes of interest can be introduced into a plant, improving a previously productive plant by the specific addition of cloned genes/traits of interest. The basic concept is extremely simple; introduce precisely defined genes into a single cell and generate a whole plant from that cell. The plant should be exactly the same as the starting material with the exception of the introduced transgenes, which should impart precise new and improved characteristics to the plant. So, how do you get a gene into a plant cell and target it to the nucleus? And, where are the cells that need to be targeted; ones that can either give rise to whole plants directly or give rise to the pollen or egg (germline) for successful transmission of the introduced DNA to progeny?

11.1.2.2. DNA Delivery. To consider DNA delivery into plant cells for the production of transgenic plants, introduced DNA must pass through the plant cell wall, cell membrane, cytoplasm, and then finally the nuclear membrane. These cell structures represent formidable barriers. The cell wall surface can be visualized as a stainless steel scouring pad, with the steel fibers representing cellulose fibers. The cell wall (especially the young cell wall) has some level of flexibility to allow cell elongation and movement, but it is a fairly rigid structure, held together with cement of pectin and other cross-linking materials. Although there are "holes" in the cell wall called "plasmodesmata," they connect the protoplasm of adjacent cells and are not open access for DNA introductions.

In order to deliver DNA into the nucleus, the cell wall must first be physically breached. Holes or breaks in the cell wall cannot be so severe so that the target cell is irreparably damaged but damage at some level must be done, to get a relatively large molecule of DNA into the cell. To complicate matters, plant cells are almost always hypertonic, which means that there is pressure pushing the cytoplasm against the cell wall, keeping plant tissues rigid. The pressure can be relieved temporarily by lowering the osmotic pressure within the plant cells (by drying the tissue or placing it on a medium containing sugars), and this has been shown to improve DNA introduction efficiencies by reducing leakage of cytoplasm from compromised cells (Vain et al. 1993).

Introduction of DNA into the cell is only part of the story as the nucleus is the desired destination in most cases. The chloroplast and mitochondria also contain genetic information and can take up and incorporate DNA, separately from the nucleus. But we will focus on nuclear transformation here. So, how does the introduced DNA get to the nucleus and what happens when it finally arrives? How does it pass through the cytoplasm, which is thick with organelles and cell machinery? With the physical methods of DNA delivery, avoidance is the best approach. It appears that the DNA is actually delivered to an area either adjacent to the nucleus or into the nucleus itself. Naked DNA (introduced DNA is almost always uncoated and unprotected as opposed to native chromosomal DNA which is specifically folded, organized, and coated with proteins) probably does not survive long outside of the nucleus. For biological methods of DNA introduction using *Agrobacterium*, the transferred DNA is coated with proteins, which protect the DNA from degradation and escort the DNA to the nucleus. Even if the introduced DNA reaches the nucleus, it is not precisely known what happens to this foreign DNA or how exactly it is incorporated into genomic DNA. It appears that the natural machinery of the cell, which repairs, modifies, and replicates DNA, is involved with sewing the foreign DNA into the genomic fabric of the target cell. Regions of native DNA are constantly being stripped of protective proteins, unfolded, accessed and reassembled. DNA must be tightly coiled and precisely ordered to fit into the nucleus but access to chromosomal DNA is needed for it to function. If foreign DNA is in the right place at the right time, it may slip into the reassembly process and become incorporated into the native DNA. Although presented here as a moderately haphazard process, foreign DNA must be precisely configured and introduced, show a necessary functionality, and appear native in order to be retained.

11.1.2.3. Target Tissue Status. For successful production of transgenic plants, specific plant cells, which have the ability to grow (differentiate) into whole plants, should be targeted. The ability of a single cell to grow into a whole plant is called "totipotency" and the cell that is naturally totipotent is the fertilized egg. Although it is probably true that all plant cells have the potential to grow into whole plants, that potential has not yet been reached for most cells. At this point in transgenic plant history, scientists can only regenerate plants from specific cell types in most plants. With a few plants, many different cell types are more easily manipulated to grow into whole plants though the tissue culture process (see Chapter 5). Successful production of genetically engineered plants is dependent on the coordination of DNA delivery with generation of a whole plant from the single cell, which is targeted for DNA introduction.

An ideal target would therefore be the unfertilized or fertilized egg, or even the pollen that gives rise to the fertilized egg. Unfortunately, these ideal targets do not appear to be responsive for almost all plants with the exception of the model plant, *Arabidopsis* (more on this later). The next most suitable target for DNA delivery might be the shoot meristem that gives rise to the aboveground parts of the plant. Although the meristem has been successfully targeted for DNA introduction, it is a complex multicellular structure, and the most appropriate target cells are located in the center of the structure, buried under quite a few cell layers. Surface cells are obviously more accessible for DNA delivery.

In the clear majority of cases, the target tissue used for production of transgenic plants consists of rapidly growing specialized plant cells, which have been induced to form whole plants. These cells should be physically accessible, actively dividing (DNA replication accelerates DNA integration

into the genome), and able to give rise to whole plants. These cells should also be resilient enough to tolerate the breach of the cell wall and membrane by the DNA, which is truly an intrusive event in the life of a plant cell.

11.1.2.4. Selection and Regeneration. Due to the nature of DNA introduction, only a small percentage of plant cells can usually be successfully targeted. The clear majority of cells therefore just get in the way. How do scientists pick out the rare cell that contains the foreign DNA? For almost all transformation efforts, selection is the key. Along with the gene of interest, another gene, encoding resistance to an antibiotic or herbicide, is introduced. The mixture of transformed and non-transformed cells is then exposed to the antibiotic or herbicide, and only those cells containing the resistance gene will survive (see Chapter 10). Selection refers to the ability of the transformed cells to proliferate in the presence of otherwise toxic selective agents. Resistance genes will encode for proteins that either detoxify a toxin or produce an alternate form of a target that is insensitive to the toxin. The most commonly used antibiotic resistance genes are neomycin phosphotransferase and hygromycin phosphotransferase which provide resistance to the antibiotics kanamycin and hygromycin, respectively. The most commonly used herbicide resistance gene is the *bar* gene (sometimes referred to as the *pat* gene), which encodes for phosphinothricin acetyl transferase. This enzyme inactivates the herbicides glufosinate and bialaphos. Selection for growth in the presence of toxic agents is the most common form of selection and is called "negative selection."

But transformed cells can be selected in other ways. "Positive selection" refers to the ability of a cell to survive by utilizing nutrient sources that are unavailable to non-transformed cells. As an example, a sugar such as mannose cannot be metabolized by most cells, unless mannose can be converted to the useful form of fructose, using a transgene that encodes phosphomannose isomerase. Cells containing this gene can grow on a medium containing mannose as the sole carbon source while the non-transformed cells will starve. A toxin is not used for selection. Another selection method utilizes genes, which allow cells to be visually identified and physically isolated from non-transformed cells. Introduction of "reporter genes" allows scientists to identify transformed cells through a unique characteristic, such as a new color or emission of fluorescence or phosphorescence. Introduction of the gene encoding the green fluorescent protein (GFP) imparts a fluorescent green color to plant cells, when viewed under high energy blue or UV light. Cells or clusters of cells containing GFP can be visually detected and physically isolated (see Chapter 10).

Once transformed cells have been recovered and purified from non-transformed cells, whole genetically engineered plants can be recovered through the tissue culture regeneration process.

11.2. *AGROBACTERIUM TUMEFACIENS*

Agrobacterium tumefaciens is a soil-born bacterium that has been rightfully called the "natural plant genetic engineer." Over its evolutionary journey, this bacterium has developed the unique ability to transfer part of its DNA into plant cells. The DNA that is transferred is called the T-DNA (for transferred DNA), and this DNA is carried on an extrachromosomal plasmid called the Ti plasmid (for tumor-inducing plasmid). For the bacterium to be used in plant transformation, scientists modified the Ti plasmid so that it no longer causes tumor formation in infected plant cells, but the T-DNA region is still transferred. As opposed to DNA transfer methods that utilize direct uptake of DNA into plant cells, the use of *Agrobacterium* seems to be more complex because two different biological systems (bacteria and the target plant cells) are involved. If this biological interaction is understood and embraced, *Agrobacterium* can provide the most reliable and consistent method for plant transformation.

With introduction of DNA without a biological vector (direct DNA uptake), it appears to be necessary to deliver the DNA to the nucleus of the target cell but with *Agrobacterium*, the T-DNA itself possesses the necessary signals for delivery there. Most direct DNA introduction systems

require expensive instrumentation, but *Agrobacterium* is simply prepared by growth on an appropriate medium and inoculated on the plant tissue. Additional claims of simpler foreign DNA insertions and more consistent transgene function in plants transformed with *Agrobacterium*, may or may not be valid, and this appears to depend more on how the DNA is delivered with direct DNA introduction systems than on any inherent problem with the method. Advances over the past 25 years in our understanding of the *Agrobacterium*-mediated DNA transfer process have led to tremendous increases in efficiency, and this natural transformation vector has become the workhorse for gene introduction in plants.

11.2.1. History of *Agrobacterium* Research

In nature, wild strains of *A. tumefaciens* cause a disease in plants called "crown gall." Crown gall disease remains a problem with many horticultural plants, notably on roses, grape, euonymus, and dahlia (Fig. 11.1). The main symptom of the disease is a gall or tumor that forms on the crown of the plant. The crown is the part of the plant that lies at the soil/air interface. This disease was a mystery to plant pathologists for many years as it does not always follow Koch's postulates, which specify that the extract from an infected organism should cause the disease when re-inoculated on a healthy plant. Also the tumors that were formed on plants would continue to grow in the absence or any microorganisms. For some time, the plant tumors were thought to be similar to some types of human cancer, but this was an incorrect assumption. Why do plant cells, infected with wild-type *Agrobacterium* grow as a tumor?

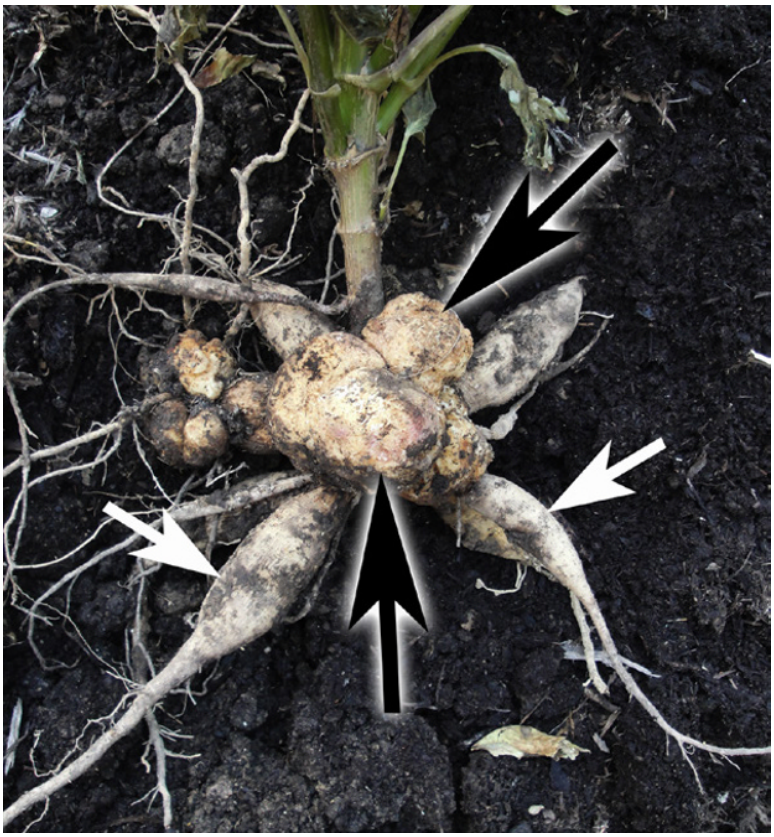


Figure 11.1. *Agrobacterium*-induced crown gall tumors (black arrows) growing next to vegetative tubers (white arrows) on the root system of dahlia.

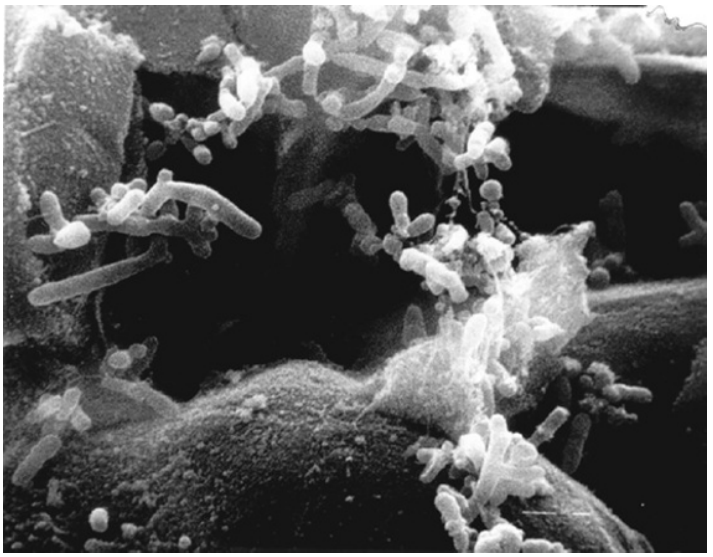


Figure 11.2. *Agrobacterium* growing on soybean tissue.

Unraveling the mystery of the disease is a fascinating story in itself and has led to the use of the bacterium for genetic engineering research. For crown gall disease, wild-type *Agrobacterium* invades wounded tissues of dicotyledonous plants. The crown is a suitable entry point as the stem is often split or torn here, as the aboveground plant moves back and forth in the wind and may rub against soil components. After entering the wounded plant, the bacteria (bacterium is singular; bacteria is plural) may either colonize dead and dying cells or simply attach themselves to the outside of a wounded living cell (Fig. 11.2). Through a series of chemical signals that are sent from the plant cell to the bacterium, *virulence* genes are activated in the bacterium that cause the bacterium to enter its virulence mode. Some of the more important mechanisms are outlined in the next section. In the end, the T-DNA is excised from the bacterium and delivered to the genome of the target plant cell.

For wild-type bacteria, the T-DNA contains only a few genes, which encode enzymes leading to the production of plant hormones and an opine, which is a nitrogen-rich organic compound that is a suitable food source for the bacterium. Tumors are formed following rapid plant cell division, as a result of hormone production in the plant cells. The opines that are produced in the tumor are used by bacteria on the tumor, within spaces in the tumor, or in the soil below the tumor. The bacteria do not colonize living, dividing tumor cells, and these tumor cells can be grown in tissue culture without added hormones. Generation and analysis of some *Agrobacterium* mutants, which contained disrupted hormone synthesis genes, helped to understand parts of the story. If one of the hormone biosynthesis genes was disrupted and the bacterium was inoculated onto tobacco plants, the tumor would produce a mass of roots. If the other hormone biosynthesis gene was disrupted, a shooty tumor would result from tobacco inoculation. If both genes were disrupted, no tumor would form (Fig. 11.3). This hormone effect was suspiciously similar to results obtained with tobacco callus in tissue culture, and these different tumor phenotypes were correctly identified as resulting from an altered hormone balance in this tissue. Much of the research that showed the transfer of DNA from the bacterium to the plant cell and even speculation on the use of this process to improve plants was put forward by the "Agrobacterium Queen," Mary-Dell Chilton (see Life Box 1.2) and her colleague Marc Van Montagu from Ghent University. Scientists from Monsanto were also central to these early discoveries and focused their efforts on developing the use of *Agrobacterium* for the improvement of crop plants.

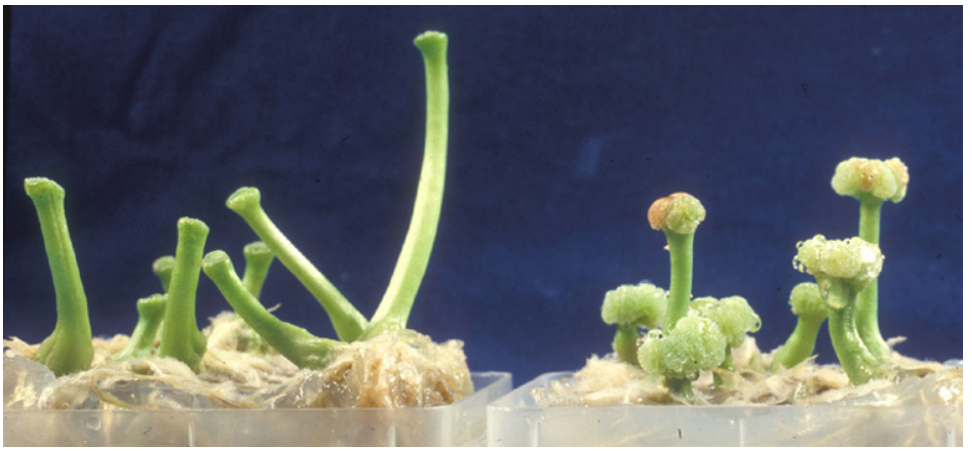


Figure 11.3. Sunflower seedling hypocotyls inoculated with *Agrobacterium* without (left) and with (right) hormone biosynthesis genes.

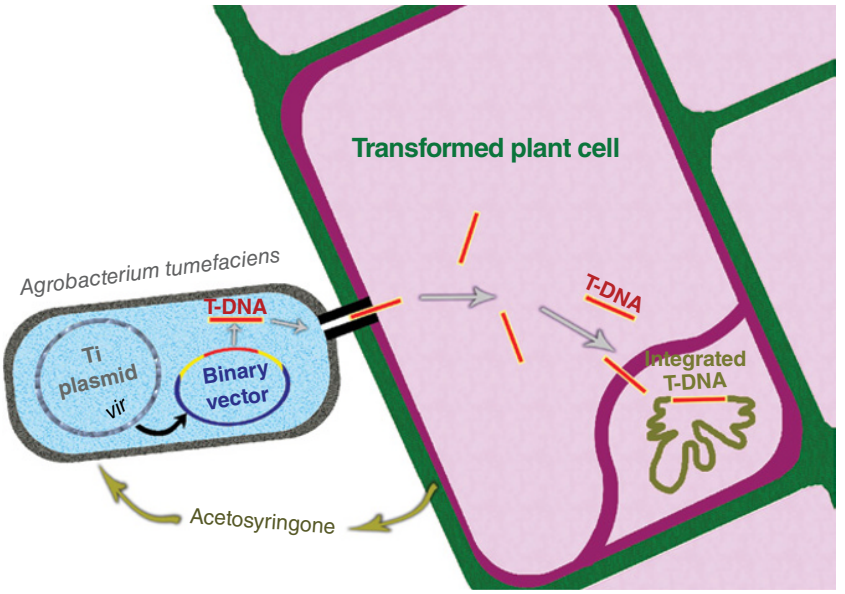


Figure 11.4. Simple schematic of *Agrobacterium*-mediated transformation of a plant cell, showing production of acetosyringone by the plant cell, induction of the *vir* genes on the Ti plasmid, generation of the T-strand from the binary vector, transport through the bacterial pilus, and integration into plant chromosomal DNA.

11.2.2. Use of the T-DNA Transfer Process for Transformation

The transition from making tumors on stems of susceptible plants to routinely transforming wheat and corn with specific genes of interest resulted from multiple advances in the understanding of both the T-DNA process and the interaction of bacteria with plant cells. Since there are many thorough reviews on the mechanism of *Agrobacterium* T-DNA transfer (Zambryski 1992; Gelvin 2012), only a few key features as relating to transformation will be presented here (Fig. 11.4).

To start, the plasmid that is used as a vector for *Agrobacterium*-mediated transformation has been whittled down to contain only the essential components. *Agrobacterium* vectors are called "binary vectors" because they are the second of two plasmids that are involved in the overall process. The

genes that provide much of the machinery for the transfer of the DNA are retained on a modified Ti plasmid with all of the T-DNA removed—the "helper plasmid." The modified Ti plasmid is still quite large. The second, small, binary vector contains primarily a modified T-DNA without the hormone and opine biosynthesis genes. The binary vector contains components that allow the plasmid to be retained in the bacterium, and the left and right "borders" of the T-DNA region. Genes of interest are cloned between the borders, which are recognition sequences for the T-DNA processing machinery. In simple terms, the T-DNA processing machinery on the modified Ti plasmid directs the processing of the T-DNA on the binary plasmid.

After *Agrobacterium* is inoculated on the appropriate plant tissue, the bacteria may recognize the target tissue as a suitable host; remember this bacterium is a pathogen that infects plant tissue. However, the strains used in plant transformation are "disarmed," meaning they have been genetically altered so that they no longer cause crown gall disease. Chemical signals are put out by both the plant tissue and the bacteria. Wounded plant tissues from appropriate plants produce acetosyringone, which activate the bacterial *virulence* (*vir*) genes, which initiates the T-DNA transfer machinery. Acetosyringone is a cell wall component derivative, which is released when the cell wall is damaged. Not all wounded plant tissues produce acetosyringone, and the absence or poor production of acetosyringone by monocot cells originally made it difficult to impossible to produce transgenic monocots using *Agrobacterium*. Addition of synthetic acetosyringone to the inoculated plant tissues allows *Agrobacterium*-mediated transformation of monocots to proceed and tremendously enhances transformation of other moderately susceptible target plants. Once the bacteria infects plant tissue, most plants will respond by trying to fight off the invasion, either by producing antipathogenic compounds or sacrificing cells adjacent to the infected region to prevent spread of the invasion. Pathogens, in turn, have developed methods to introduce regulatory compounds into plant cells, in an attempt to shut down the defensive machinery of the target cell. Although some of these mechanisms are known, some are still being investigated and a more thorough understanding of the infection process will allow further increases in the efficiency of *Agrobacterium*-mediated transformation.

Once the *vir* genes are activated, the T-DNA on the binary vector is processed for transport to the target plant cell. Some of the *vir* gene products excise the T-DNA from the binary plasmid as a single-stranded DNA molecule, while other *vir* gene products coat the T-DNA to prevent degradation. Yet additional *vir* gene products bind to the T-DNA to act as navigators or signals to direct the DNA out of the bacterium, through the plant cytoplasm, and to the nucleus. Through the action of other *vir* genes, the bacterium produces a pillus, which is the conduit for transfer of the T-strand (the single-stranded, coated, signal containing T-DNA is now called the "T-strand") from the bacterium to the target plant cell. The pillus is essentially a protein tube, which extends from the bacterium through the cell wall and into the cytoplasm of the target cell. After the T-strand is delivered to the nucleus, the last role of the signal protein on the T-strand is to find and nick the host DNA as an insertion point for the T-DNA. The T-DNA appears to insert primarily into transcriptionally active regions of DNA, as regions that are actively transcribing, are more exposed and accessible.

11.2.3. Optimizing Delivery and Broadening the Taxonomical Range of Targets

As more is learned about the mechanisms underlying *Agrobacterium*-mediated transformation of plant cells, the efficiency of the process will undoubtedly increase. The four main approaches for improving transformation are the following: (a) increase delivery of the bacteria, (b) induce the *vir* genes, (c) minimize defense responses of the target tissue, and (d) select compatible plant cultivars and *Agrobacterium* strains.

Numerous methods have been developed to increase the delivery of the bacteria to the target plant tissue. Since the bacteria infect though wounded tissues, and wounded tissues generally produce acetosyringone, most of these methods strive to either increase overall wounding or call for precision wounding. The most common tool for wounding of the target tissue is the scalpel, which is simply used to excise plant tissues. When the tissue is cut, this presents a suitable binding/entry point for

the bacterium. Wounding can be increased by scoring the target tissue multiple times, with a scalpel blade. Severe wounding of this sort will eventually lead to a loss of the ability of the plant tissue to regenerate. Precision wounding using either sonication or particle bombardment (later in this chapter) results in the generation of large numbers of extremely small wounds. Precision micro-wounding, if done properly, does not extensively damage the tissue structure and tremendously increases the number of entry points and attachment sites for the bacteria.

Induction of the *vir* genes through the addition of acetosyringone has led to routine transformation of plants that were initially not thought to be susceptible to *Agrobacterium*-mediated transformation. Although acetosyringone may not improve transformation of very susceptible plant species, such as tobacco (they already produce sufficient levels), it is routinely added during the co-culture period for most other plants. Co-culture is the time period when bacteria are permitted to invade, infect and transform plant cells. The end of the co-culture period occurs when appropriate antibiotics are added, to eliminate the bacteria after their job is done. Results, similar to acetosyringone addition, can be obtained with the use of *vir* gene mutants, which were modified so that the *vir* genes are always active, even in the absence of acetosyringone.

The area that may hold the most promise for future increases in efficiency of *Agrobacterium*-mediated transformation is the alteration of response of infected plant cells to the bacterium. During the interaction between *Agrobacterium* and plant cells, elevated peroxidase activity and subsequent oxidation may cause tissue browning and cell death. Improvements have been made in transformation frequencies following the addition of reducing agents, which minimize the effects of oxidizing agents produced by infected plant tissues. The most commonly used agents are cysteine, dithiothreitol, silver nitrate, and ascorbic acid. In addition to reducing agents, enormous potential exists for using agents and genes that eliminate or reduce programmed cell death (PCD) in target tissues. Although PCD is a good natural defense mechanism for sequestering or localizing an infection and preventing spread by death of the infected cells, a reversal of this defense leads to higher transformation efficiency.

11.2.4. Strain and Cultivar Compatibility

Earlier in this chapter, the greater susceptibility of dicots to *Agrobacterium* was emphasized over monocots. Ingo Potrykus, the inventor of Golden Rice and one of the greatest plant biotechnology scientists of all times (see Life Box 9.2), even concluded in the early days of transformation that it was "probably impossible to transform cereal plants with *Agrobacterium*" (Potrykus 1991). Well, this is now routine. Although cereal transformation issues were largely overcome by adding acetosyringone, additional improvements were made by using the right strains of *Agrobacterium* with compatible plant cultivars. Because *Agrobacterium* is a plant pathogen, certain bacterial strains are more virulent on some plants and even some cultivars, within species. If it not critical to introduce a gene into a specific plant cultivar or variety using *Agrobacterium*, a number of varieties should first be evaluated with a number of *Agrobacterium* strains, to check for compatibility, which means higher transformation rates. In most cases, select plant cultivars are often used, because they are both responsive to *Agrobacterium*-mediated transformation and can be more easily regenerated from tissue culture.

The problem with the previous suggestion, to evaluate different *Agrobacterium* strains with different plant varieties, is the scarcity of laboratory strains of *Agrobacterium*. The most common *Agrobacterium* strain, used by transformation scientists all over the planet, is EHA105 (named after Elizabeth Hood, who is recognized for disarming the strain (Hood et al. 1993)). EHA105 is a derivative of C58, which was originally isolated from a gall from a cherry tree (the "C" in C58 stands for "cherry"). Although this EHA105 strain has been used for gene introduction in cherry, it has been mostly used for transformation of other plants. It may make sense to isolate additional *Agrobacterium* strains, from tumors or soil samples from plants or the fields where certain plants grow, to find novel strains that may coexist with the plants in the field. These new strains can show

tremendous enhancements in transformation efficiency of certain plants (Benzle et al. 2015). Strain isolation can be challenging, as *Agrobacterium* strains have to be purified from other bacteria in the samples. Once a suitable *Agrobacterium* strain is isolated and identified, disarming of that strain is required in order for it to be used in plant transformation. Disarming is easier now than it used to be, because of the currently available genomics tools. But, this is still an effort. Considering that *Agrobacterium*-mediated transformation of plants is based on a biological plant/pest compatibility and most transformation laboratories use only a handful or less of *Agrobacterium* strains, tremendous potential exists for improving transformation efficiencies, if novel strains can be identified.

To take this concept further, could it be that *Agrobacterium* is not that unique? Maybe. There are other bacteria or pathogens that have the ability to modify their host by transferring parts of their DNA to plant cells. Viruses can take over their host through introduction and expression of their genetic material. But that introduced material does not often integrate into the host genome and represents a temporary takeover. However, other bacteria may share *Agrobacterium*'s unique talent. In an effort to work around the intellectual property restrictions currently in place for using *Agrobacterium* for plant transformation, *Ensifer adhaerens* was identified as another related bacterium that also can transfer DNA to plant cells, resulting in permanent genetic changes (Rudder et al. 2014). Although this bacterium is currently being evaluated by a number of different transformation laboratories as an alternate organism to transfer the T-DNA from *Agrobacterium*, it is unclear if this approach can be widely repeated or if it has the same range of targets that exists for *Agrobacterium*.

11.2.5. Agroinfiltration

Certain situations exist where rapid manipulation of gene expression is needed, but it is not necessary to transform a cell and take the time to recover a whole transformed plant. Why? In some cases, the effects of introducing a new gene or lowering the levels of expression of a native gene can be very quickly determined using "agroinfiltration." For agroinfiltration (Vaucheret 1994), *Agrobacterium* is injected or infiltrated into leaves of a suitable target plant. Although *Nicotiana benthamiana* is a model for this approach, it has been successfully applied to a number of different plants, where rapid evaluation of gene expression is desirable (King et al. 2015). For this method, an *Agrobacterium* suspension is forced into the internal leaf air space by tightly holding a syringe (without the needle) to the leaf and pushing the plunger. A variation of this method requires dipping the plant into an *Agrobacterium* suspension to wet the leaves and then applying vacuum to force the bacterium into the internal leaf air space. Alternately, the plant can be submerged in an *Agrobacterium* suspension and a vacuum applied directly. To enhance the levels of gene delivery and spread, the T-DNA can be modified to contain viral gene components to launch the viral amplification and transfer machinery, making this method very efficient for production of transgene product (Fig. 11.5) in plants without transfer to the next generation.

11.2.6. *Arabidopsis* Floral Dip (Clough and Bent 1998)

Arabidopsis thaliana remains the most prominent model species for plant genomics. The genome and the plant itself are small, the generation time is rapid, and it is ridiculously easy to transform. The floral dip method was developed for *Agrobacterium*-mediated transformation of *Arabidopsis*, and no other plant currently responds similarly, even after large efforts to apply this approach to other plants. The floral dip methods results in generation of independently transformed seeds, as a result of *Agrobacterium*-mediated transformation of the female gametophyte or the egg.

For floral dip, *Arabidopsis* plants are simply immersed in a suspension of *Agrobacterium*. After the dipping treatment, plants are maintained under high humidity for a few days and allowed to eventually flower, typically by selfing, and set seed. Since *Arabidopsis* produces so many seed and the plants are so small, seeds can be easily planted on selective media or seedlings/plants can be screened for a certain characteristic or phenotype to recover whole transgenic plants.

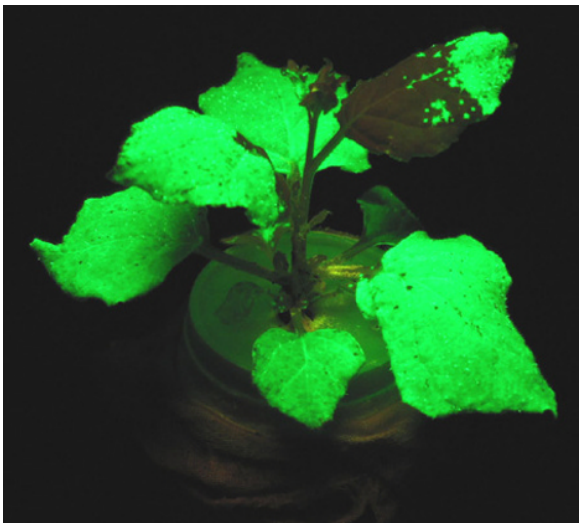


Figure 11.5. Agroinfiltrated *Nicotiana benthamiana* plants showing high levels of GFP expression. The aerial parts of the tobacco plant were submerged in an *Agrobacterium* suspension and the plant was then placed under vacuum for infiltration. Courtesy of John Lindbo. (*See insert for color representation of the figure.*)

It appears that the *Agrobacterium* proliferates or multiplies at low levels and coexists within the tissues of the plant. They do not invade the cells of the plant as most other pathogens do, but they bind to suitable target cells for DNA delivery. If *Agrobacterium* transforms a somatic or vegetative cell within the plant, this is probably a terminal event. In the plant, a transformed leaf or stem cell will not give rise to anything other than another leaf or stem cell. For recovery of whole transgenic plants, the goal is targeting germline cells: cells that will contribute to the fertilized egg. For *Agrobacterium*-mediated transformation of *Arabidopsis*, the egg appears to be the serendipitous target of T-DNA delivery, leading to the production of transgenic seed. Usually each seed is from a different transformation event.

Why is *Arabidopsis* so easy to transform? Why don't corn, wheat, rice, and soybeans work the same way (soybean transformation has been very inefficient for the past 20 years). As we learn more about the transformation process, it may eventually be possible to recover transgenic corn, wheat, rice, and soybean with the same ease as transgenic *Arabidopsis*. Transformation efficiencies of these plants have increased tremendously over the years, but it is still tough. For now, inefficiencies in transformation remain a mystery and a reason for transformation scientists to continue working on discovering improvements in strains and protocols.

11.3. PARTICLE BOMBARDMENT

11.3.1. History of Particle Bombardment

Particle bombardment refers to a method where heavy metal particles (~1 μm gold or tungsten) are coated with DNA, accelerated toward the target tissue, and penetrate the cell wall to rest either adjacent to or directly in the nucleus. The DNA on the particles somehow finds its way to the native DNA of the target cell, where it becomes integrated into the chromosome to become a permanent addition to the genome.

The term "particle bombardment" can be used interchangeably with the similar terms "microprojectile bombardment," "biolistics," "particle acceleration," and "gene gun technology." The term that is currently most often used is "particle bombardment."

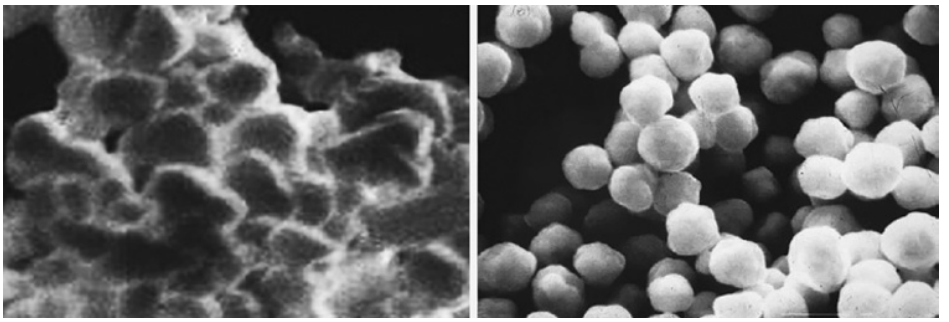


Figure 11.6. Tungsten (left) and gold particles (right) used for particle bombardment, prior to DNA precipitation. Gold particles are more uniform and spherical than tungsten particles.

As opposed to *Agrobacterium*, which is a biological vector for DNA introduction, particle bombardment is a purely physical method for DNA delivery. The DNA is physically precipitated onto metal particles and those particles are then rapidly accelerated toward the target tissue using a device: a gene gun. The particles penetrate through the cell wall by punching holes in that rigid structure, and they continue until they are stopped by the density of the target tissue. To visualize the process, imagine bullets or shotgun pellets, penetrating a thin piece of wood to enter the water beneath the wood. The wood represents the cell wall and slows down the particles abruptly, while the water gradually slows them down further until they stop. The reference to bullets is no coincidence.

Particle bombardment was invented by John Sanford and colleagues at Cornell University in the mid-1980s. The approach was further developed and optimized by Ted Klein (see Life Box 11.1), a postdoc in John Sanford's laboratory. Conceptually, a 22 caliber rifle, loaded with blanks, was first used to evaluate the damage to plant tissue from the shock wave resulting from an ignited powder load. The "gun" with "bullets" concept was further perpetuated with the introduction of the first commercial device, which used a 22 caliber powder load to generate a controlled explosion to accelerate small tungsten particles down the barrel of a modified gun. Between "shots," the particle bombardment device had to be cleaned with a gun-cleaning swabs and brushes. Later versions of particle guns used other types of forces to generate the energy required to accelerate the small particles. The required violent forces, needed to accelerate the particles, could be created by generating high voltage arcs across a gap or by using high pressure air or CO_2. It was not unusual at the time to perform bombardments with muffling headsets, to dampen the sound from the early devices. Today, in most laboratories, high-pressure helium is used to generate the force needed to accelerate small gold particles (Fig. 11.6) toward the target tissue. Helium is preferred since it is inert and has a high expansion coefficient, meaning that it can be compressed and expands rapidly when released into the air or a vacuum. DNA is first precipitated onto the particles, which are then placed as a monolayer on a mylar carrier sheet, called a "macrocarrier." The term "macrocarrier" refers to the structure, which carries the particles, while the term "microcarrier" was originally designated for the particles (as they are small and carry the DNA). The controlled explosion, used to accelerate the macrocarrier, is provided by high-pressure helium, which is released from a small chamber following the breakage of a rupture disc, designed to break at specific pressures. The macrocarrier, with the particles on one side, travels a short distance and smashes into a screen (looks like window screen), stopping the macrocarrier and allowing the particles to continue along their path. In most cases, the whole procedure is performed under partial vacuum because the presence of air slows down the particles. A partial vacuum, applied for a short duration, does not appear to damage the biological targets.

Although there are numerous versions of particle bombardment devices (Fig. 11.7), they all utilize the same basic approach. The main manufacturer of particle bombardment devices is BioRad,



Figure 11.7. Two different particle bombardment devices: The commercially available PDS1000 He (BioRad) (left) and the non-commercial Particle Inflow Gun (right).

who offers two different versions of the device. One version is a large, heavy, vacuum-utilizing unit that sits on a lab bench, whereas the other version is handheld and moderately portable. The handheld "Helios" device has received more attention for gene therapy work while the large, bench-top unit is standard in many plant transformation laboratories. There is certainly a cost associated with all of these devices, which limits the use of particle bombardment for those laboratories with insufficient resources.

11.3.2. The Fate of the Introduced DNA into Plant Cells

For DNA introduction using particle bombardment, DNA is first precipitated onto the particles using either calcium chloride or ethanol, which are commonly used for DNA precipitation. When the DNA precipitates, it sticks to whatever is at hand. It is unclear how "tightly" the DNA is bound to the particles, but it must be able to withstand the incredible force of acceleration and cell wall/cytoplasm penetration and also come off the particles after delivery.

During bombardment, the majority of the metal particles do not find their target. Most of the particles either embed in the cell wall, enter the vacuole, or end up in somewhere else in the cytoplasm; only a few reach the nucleus. After all, thousands of particles are delivered using literally a "shotgun" approach. Evaluation of those cells that express the introduced DNA shows that the overwhelming majority of those cells (>95%) have particles in the nucleus (Yamashita et al. 1991). It is unclear how many cells contain particles but do not express the introduced DNA. Unlike *Agrobacterium* where integration of the introduced DNA into the plant chromosomal DNA is orchestrated by bacterial proteins that are bound to the T-strand, particle bombardment results in the introduction of naked DNA. Clues to the fate of the introduced DNA can be taken from studying the final arrangement of the integrated DNA within plant chromosomal DNA.

In general, the patterns of DNA integration in the plant chromosome resulting from particle bombardment are very complex. To be more specific, it can be a real mess. Usually, the introduced DNA integrates into a single site (locus) on the chromosomal DNA. However, the introduced DNA can also integrate at multiple sites, which makes analysis more difficult. To complicate the situation

further, it is common to obtain multiple copies of the transgene in each integration site. And it gets worse. The copies can be partial copies, with varying orientations. In addition (last thing), the introduced DNA appears to be mixed or interspersed with plant genomic DNA (Pawlowski and Somers 1998). Imagine the replication and repair machinery of the nucleus as an army of overworked, frantic, multiarmed, DNA tailors. The DNA tailors are supposed to make exact copies of chromosomal DNA and fix any small mistakes, while they are sewing huge amounts of new DNA strands. They are working fine until the whoosh of this huge boulder (1 μm particle) overhead which is carrying DNA. It looks like plant DNA, so they take what they can and use it, in their sewing operation. It is not a perfect fit, but they are frantic and under time constraints to get the entire chromosomal DNA replicated before the cell divides. For particle bombardment, it is unclear if the particles actually physically break the chromosomal DNA or just deposit DNA in the proximity of the replicating parts of chromosomes. It is clear that the introduced DNA can integrate into chromosomal DNA, with very complex patterns. But, complex integration patterns can be largely controlled by manipulating the configuration of the introduced DNA (see Sections 11.3.3 and 11.3.4).

11.3.3. The Power and Problems of Direct DNA Introduction

As particle bombardment is a physical method for DNA introduction, complications from biological interactions with the plant (as with *Agrobacterium*-mediated transformation) are avoided. A wide variety of plant tissues can be used as targets for particle bombardment. These range from embryos, seedlings, shoot apices, leaf discs, microspores, and immature pollen grains, to potato tubers and nodes (Altpeter et al. 2005). Although the foreign DNA integration patterns (discussed earlier) can be very complex, this mechanism for DNA recombination and integration can be an advantage. Various DNAs can be mixed and co-introduced: a method called "co-transformation." From 12 to 15 different DNAs have been successfully co-transformed into soybean (Hadi et al. 1996) and rice (Chen et al. 1998). This complex integration of transgenes is potentially useful for pathway engineering, where it is necessary to introduce multiple genes at once.

Particle bombardment remains the main method used for transformation of chloroplasts and mitochondria. Plastid transformation (Maliga 2014) is useful in cases where large amounts of the transgene product are needed. The integration of foreign DNA into plastid DNA is also simple because integration events are less complex, compared to nuclear transformation. For plastid transformation, the foreign DNA is able to be targeted precisely into sites that have similar sequences in the desired plastid DNA locus, using "homologous recombination." Another advantage of plastid transformation is that, in most plants, plastids are maternally inherited, thus avoiding the possibility of long distance transgene spread via pollen. But, like the floral dip method, this technique is currently limited to a small number of species.

In hand with the numerous merits of particle bombardment, there are certain drawbacks that limit its use. The main perceived limitations are the randomness of DNA integration and the high copy number of introduced DNAs. As with most methods of DNA introduction, the position and orientation of the transgene in the plant chromosome will differ with every transformation event. The location of the transgene within the target chromosome will influence the expression of that gene. Transgenes in more transcriptionally active regions of genomic DNA will express at higher levels, while integration in less active areas will lead to lower expression. These are called "position effects," in which transgene expression is influenced by its position within the genome. More importantly, the number of copies of introduced DNA can be very high, leading to inactivity of the introduced DNAs (Taylor and Fauquet 2002). If a company wants to commercialize a transgenic crop, an event with one integration site is almost always desired. One might think that the presence of many copies of a particular transgene would result in very high expression. But, expression of the transgene is often downregulated by the plant, a phenomenon known as "co-suppression," "homology-dependent silencing," "RNA interference (RNAi)," or "RNA silencing." Selection of plant cells/tissue showing uniform transgene expression is critical. Several techniques have been developed to minimize

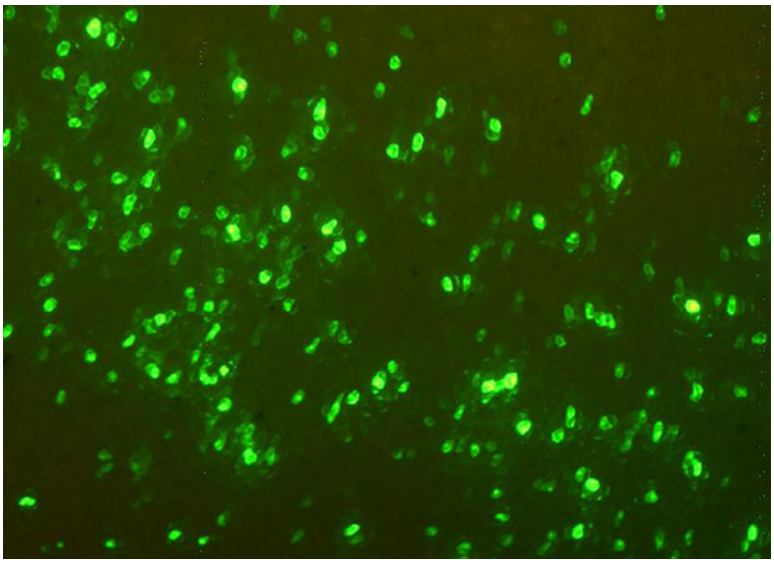


Figure 11.8. Particle bombardment-mediated transient GFP expression in lima bean cotyledonary tissues. This target tissue is flat, non-pigmented, and ideally suited for tracking GFP expression in individual transiently transformed cells. (*See insert for color representation of the figure.*)

variation in transgene expression from particle bombardment. These methods are similarly applicable to other direct DNA introduction methods (later in this chapter).

11.3.4. Improvements in the Control of Transgene Expression

Variation in transgene expression resulting from particle bombardment can be reduced to some extent by modifying the introduced DNA. Since high copy number integration appears to lead to transgene silencing, any method of controlling copy number could lead to an improvement in consistency of transgene expression. To start, reducing the concentration of DNA on particles that are shot into cells appears to reduce the copy number of the transgene in the target cell. High concentrations of DNA are still used in many cases and are a remnant of early optimization strategies. The use of high concentrations of DNA results in high levels of "transient expression," which is used to optimize DNA delivery conditions. Transient expression refers to a rapid increase followed by a decline in expression of transgenes, which can measured and quantified (Figs. 11.8 and 11.9). As a result of these levels of high transient expression, DNA concentrations, which are much higher than necessary, are often used for stable transformation studies. The beneficial effects of lower concentrations of DNA on stability of transgene expression should be evaluated for each different target tissue. Copy number of the introduced transgene can also be lowered by simplifying the form of the introduced DNA. Simple integration patterns result if a fragment of DNA containing only the gene of interest is used. When the backbone of the cloning plasmid is eliminated from the bombardment precipitation mix, this results in low-copy transgene integration (Agrawal et al. 2005).

11.4. OTHER METHODS OF TRANSFORMATION

11.4.1. The Need for Additional Technologies

With the two main methods for DNA introduction, why are additional methods needed? Isn't this enough? In the scientific community (and for humanity in general), the theme is "bigger, better, stronger, and faster." Certainly, plant transformation is achievable and transgenic plants have been

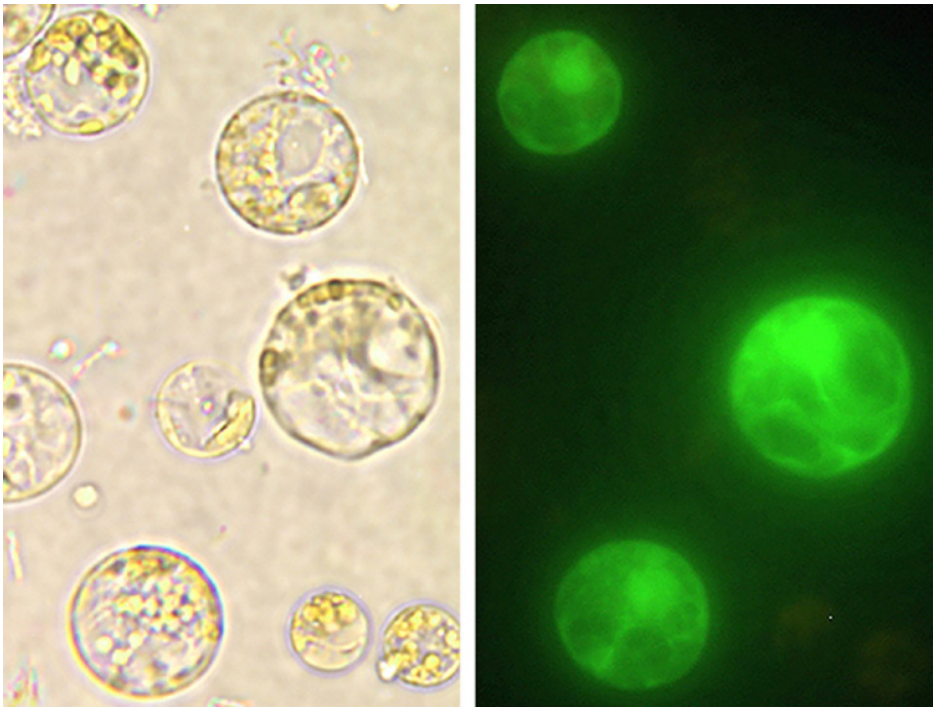


Figure 11.9. Maize protoplasts, electroporated with a *gfp* gene, showing bright field (left) and with GFP filters (right). Courtesy of JC Jang. (*See insert for color representation of the figure.*)

obtained using all of the plants of major economic importance. But efficiencies of existing methods can always be increased, and new methods may yield even higher transformation rates. Floral dip of *Arabidopsis* is very straightforward and efficient, but further improvements could be made in the recovery of more transgenics and, more importantly, application of this method to other plants would be quite useful. In addition, most of the methods that have been presented, including those that will be presented later, are protected by patents. The status of intellectual property drives much of plant biotechnology, and the methods for transgenic plant production are no exception. New transformation technologies will probably be protected by patents but the availability of more choices is always beneficial.

The additional technologies presented here do not represent a complete or thorough list. The methodologies are presented to provide a sampling of the types of ideas that have been generated since the dawn of transgenic plant production in the mid-1980s.

11.4.2. Protoplasts

As presented earlier in this chapter, the cell wall represents the greatest barrier to introduction of DNA into plant cells. When the plant cell wall is enzymatically or physically removed, protoplasts (Fig. 11.9) are the end result. Protoplasts are very fragile single cells, which must be maintained in an osmotically and nutritionally balanced medium to prevent lysis. They are typically generated using enzyme mixtures of cellulases and pectinases to digest cell walls, and mannitol is often used to maintain the osmotic integrity of these naked cells. Protoplasts can be generated from many different types of tissue, but young leaf mesophyll tissues and embryogenic cultures are the most common. Although protoplasts can be manipulated in a number of ways in the laboratory, they are most often used either for DNA introduction or to generate fusion hybrids.

The two main methods used to introduce DNA into protoplasts are electroporation and polyethylene glycol (PEG) treatment. For electroporation, protoplasts are placed in a DNA solution between two electrodes and exposed to brief pulses of high voltage current. The pulses cause pores to form in the membrane and the DNA then enters the cells. PEG treatments are also performed in the presence of DNA and probably also result in the formation of pores, from membrane destabilization.

With their cell walls removed, protoplasts can be manipulated in additional ways that are not possible with intact plant cells. DNA can also be introduced into protoplasts using microinjection, which is the most common method used for transformation of animal cells. Microinjection utilizes precisely drawn and cut glass needles, which will shatter if pushed into an intact plant cell. Surprisingly, protoplasts can also be very efficiently transformed using *Agrobacterium*. The bacteria are able to very effectively adhere to and transform protoplasts as the protoplasts are regenerating new cell walls.

Although protoplast transformation can be extremely efficient with greater than 50% of the cells receiving DNA, tremendous problems are encountered when attempting to recover whole plants from these single cells. Whole transgenic plants have been recovered from a variety of plants using protoplast transformation, but it is not very often used today for generation of transgenic plants. Because DNA introduction efficiency can be very high, transient expression in protoplasts is routinely used for analysis of factors that modulate gene expression (Sheen 2001).

11.4.3. Whole Tissue Electroporation

Although electroporation can be used for very efficient transformation of protoplasts, application of electric pulses to whole tissues can also result in DNA introduction, although at reduced rates of efficiency (D'Halluin et al. 1992). With the cell wall intact, the formation of pores in the cell membrane is of limited value for DNA introduction. Whole tissue electroporation has been successfully used with rapidly growing tissues which contain thin, newly formed cell walls. Partial enzymatic digestion of whole tissues using cellulases and pectinases can remove enough of the cell wall to allow DNA introduction using electroporation of partially "intact" tissues.

11.4.4. Silicon Carbide Whiskers

Developed originally for DNA introduction into insect eggs, use of silicon carbide whiskers have been successfully applied for DNA introduction into plant cells (Kaeppler et al. 1990). Silicon carbide whiskers are long rigid two-pointed microscopic "spears," which are added to plant cells and DNA and then vortexed. The spears or whiskers are approximately 1 μm thick and 15–50 μm long. Although the analogy of "being in a Jacuzzi with porcupines" has been used to describe this technology, the shaking motion is much more violent and is probably more closely akin to a paint mixer found in hardware stores. It seems that the whiskers enter the cell with DNA as they are trapped between two cell clusters as they collide. The low efficiency of transformation using silicon carbide whiskers along with disposal under conditions similar to asbestos makes this method unsuitable for most laboratories.

11.4.5. Viral Vectors

Since most plants can be infected by numerous viruses, viral vectors could potentially be used as another "natural" DNA introduction method for plants. Using their own transport mechanism, viruses can spread on their own throughout their host, so introduction of a virus into a single cell can eventually lead to the presence of virus genes in many cells of the inoculated plant. Although viral vectors can be used for extremely efficient introduction and transport of virus genes, these genes do not integrate into the genome of the host cell. Therefore, they will not be transmitted to the next generation through the pollen and egg.

But, inoculation of viruses into plant cells can be as simple as rubbing the leaf in the presence of the virus and a single site of inoculation can lead to expression of viral genes in most of the cells of the plant (which is similar to production of a transgenic plant but is not quite the same). For successful introduction and expression, the gene of interest must be appropriately packaged in the viral genome, which tends to be less cooperative in accepting foreign DNA, especially large pieces of DNA. Viral vectors are useful for very rapid production of proteins in plants without the need to generate a whole plant from a single, transformed cell.

11.4.6. Laser Micropuncture

For direct DNA introduction into plant cells, the use of micro lasers continues with the theme of creating holes in the cell wall (Badr et al. 2005) for DNA delivery. This is perhaps one of the more elegant and least often utilized methods for DNA introduction into plant cells. Lasers are very precise in targeting certain cells, but the instrumentation required for this method is quite involved and the number of cells which are targeted is very small. As a comparison, for particle bombardment, the number of cells that transiently express an introduced transgene can be greater than 10,000 per shot. Many more cells are actually targeted—this is the number of cells that receive the DNA close to or in the nucleus and transiently express the introduced DNA. For laser micropuncture (and protoplast microinjection, earlier), cells are targeted, one at a time. It is doubtful that the use of micro lasers for DNA introduction will increase tremendously, but it is a noteworthy method for DNA introduction into plant cells.

11.4.7. Nanofiber Arrays

Successful use of nanofiber arrays (Melechko et al. 2005) for DNA introduction into plant cells has not yet been consistently obtained, but convincing results have been demonstrated using animal cells (McKnight et al. 2003). Nanofiber arrays can best be described as a microscopic "bed of nails" (Fig. 11.10). Early attempts to generate nano arrays resulted in the formation of nano pyramid-shaped structures on a silicon chip (Hashmi et al. 1995). In this early work, the surface of the chips was precisely etched away, to leave the nano pyramids. Next-generation arrays are composed of

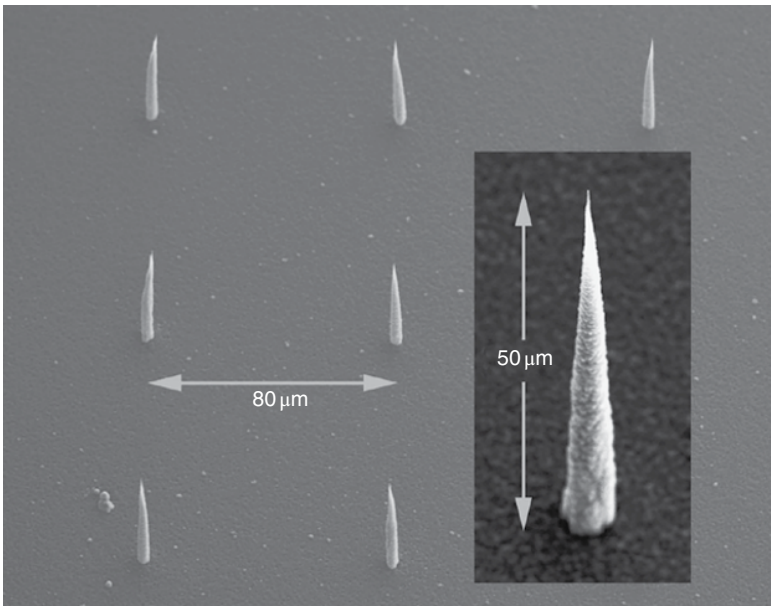


Figure 11.10. Nanofiber array with single fiber at higher magnification (inset). Courtesy of by Tim McKnight.

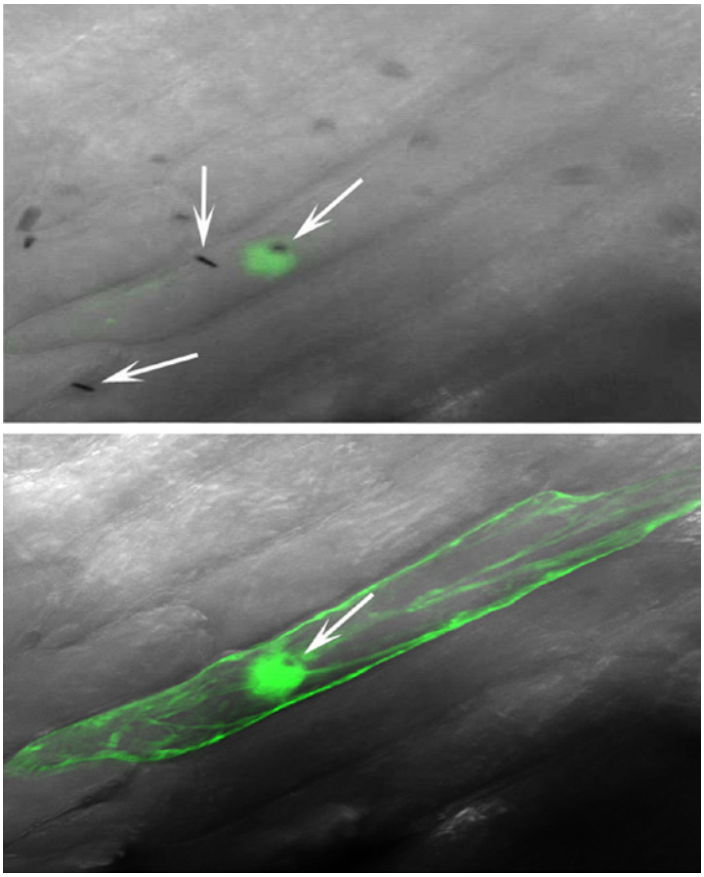


Figure 11.11. Confocal microscopy sections of onion cells showing dislodged nanofibers (upper panel, white arrows) and expression of the green fluorescent protein gene from a dislodged fiber in the nucleus (lower panel, white arrow).

long, thin structures that hold much more promise for success with DNA introduction into plant cells. Nanofiber arrays are actually grown on chips, with very precise composition, height, and spacing possible. DNA can be chemically bound to the fiber or simply precipitated onto it. For the successful DNA introduction into animal cells (McKnight et al. 2003), the arrays were stationary and the animal cells were propelled toward the chip. Cells were then allowed to grow, while still impregnated with fibers, on the chip. Although the cell wall is certainly much more of a barrier than the animal cell membrane, the fibers are sufficiently strong and rigid to allow them to penetrate the plant cell wall. And because the chip surface is covered with fibers, many cells can be targeted using a single chip. Results with onion epidermal cells show the utility of this approach (Chiera and McKnight unpublished observations; Fig. 11.11), but the high efficiency delivery of DNA-coated nanofibers directly to the nucleus of multiple plant cells remains a challenge.

11.5. THE RUSH TO PUBLISH

11.5.1. Controversial Reports of Plant Transformation

Have you ever heard anyone say, either seriously or as a joke, “I read it online so it must be true”? Anyone can display any information using an incredible variety of delivery strategies. As a result of this situation, we are now becoming more cautious of the information that we collect from personal

webpages, websites, newspapers, and local and national broadcast news. The scientific literature sets a high bar for disclosing valid scientific information to the international scientific community through the use of the peer-review process. Scientific articles are usually reviewed by two or more science experts, who will either accept the science, reject the results, or call for more experiments to be performed because the results or conclusions are unclear. It is a good system, but when the pressure to publish research articles is high and there are so many different journals to submit for publication, sometimes research results are published that should have first been more critically evaluated and tested.

In plant transformation, breakthrough technologies are highly prized and quite valuable. In addition to the notoriety that comes along with new discoveries, patent protection can provide a reasonable source of additional income, at least for university scientists. Truly new ideas in the sciences are actually not very common and most of the advances that are reported in the scientific literature represent improvements in pre-existing technology or small steps in our understanding of processes. When something really new does come along, it should be very critically evaluated, especially since there is a premium on being the first paper on a breakthrough technology. Unfortunately, there are numerous reports in the literature that initially cause quite a stir and then disappear because they did not work or worked with such a low efficiency, that they are impractical. In science, published methods should be able to be replicated in other labs. In science, it is not about who's right but what's right. At the risk of alienating other scientists, some controversial reports of plant transformation methods are listed in the text.

11.5.1.1. DNA Uptake in Pollen (Hess 1980). For one of the first reports of plant transformation, pollen from a white-flowering petunia was soaked in DNA extracted from a red-flowering petunia. When this soaked pollen was used to pollinate the white-flowering petunia, some of the resulting seeds produced plants with either partially or fully red flowers. The author concluded that the DNA must have been taken up by pollen and passed onto the seedling from the fertilization process. The authors were cautious about the interpretation of their work and came to their conclusions of genetic transformation as the most probable explanation of their results. They did the appropriate controls and noted that a small amount of red pigmentation could occur in white flowers at certain times of the year and in response to various stresses. Since that work was published over 25 years ago, no one has been able to repeat this work, after extensive efforts. The tools to test for the presence of foreign DNA were not in existence at the time this work was done and the red flower color was the only evidence for transformation. The most plausible explanation for these results is pollen contamination, which the author discounted as they had never observed this with any of their controls.

11.5.1.2. Agrobacterium-Mediated Transformation of Maize Seedlings (Graves and Goldman 1986). Although *Agrobacterium*-mediated transformation of maize is now fairly routine, this early report of *Agrobacterium*-mediated transformation of maize tissues remains quite controversial. At the time of this report, there were a few claims of *Agrobacterium*-mediated transformation of monocots and no reports for the economically important cereals. In addition to a scientific publication, this work led to the issuance of numerous patents. We must keep in mind that patents are not scientific publications and undergo a very different review process. In this 1986 publication by Graves and Goldman, maize seedlings were wounded and inoculated with *Agrobacterium*. Although transgenic plants were not recovered, the authors reported that the seedlings tested positive for the presence of opine synthase enzymes. These specific enzymes can only be produced after successful T-DNA transfer, and opine synthase analysis was one of the only tests for successful transformation at the time. This work was done before the optimization treatments, which were described earlier in this chapter, were even known. And the transformation efficiency in this paper was 60–80%, which is high even by today's standards. If copies of this paper are inspected, inconsistencies cannot be detected. However, if the original paper is carefully examined, one can see the differences between the control and experimental treatments disappear when the images showing

opine production are lightened or darkened to provide digitally equalized background levels. This paper is continuously referenced in the transformation literature, but it should not be so.

11.5.1.3. Pollen Tube Pathway (Luo and Wu 1988). The pollen tube pathway method for transformation is different from pollen transformation (see earlier text) as the pollen is not transformed, but the pollen tube is used as a vehicle for the delivery of DNA to the egg or fertilized egg. The basis of this method is the inoculation of DNA into the hollow pollen tube, where it finds its way to the freshly fertilized egg for incorporation into the DNA of the young zygote. Timing was reported to be critical as the pollen is first placed on the stigma for germination. After the pollen tube grows down the style to the ovary, the stigma is severed, leaving a narrow hollow pollen tube as an open pathway to the fertilized egg. DNA is then inoculated onto the open pollen tube where it was believed that capillary action drew the DNA in solution to the zygote. On the surface, this method appears to have some merit, but the pollen nuclei are the only things to enter the egg and the fertilized egg or zygote has the same barriers as any other young plant cell, notably the cell wall. From the 1980s through even today, there are many additional reports in the literature of the successful use of the pollen tube pathway for many different crops; almost all of these reports originate from China. Although it is very difficult to publish negative results, Shou et al. (2002) performed a very extensive study of the pollen tube pathway method in soybean and concluded that it was not reproducible. It appears that the pollen tube pathway method for DNA introduction is not feasible. In the first published report (Luo and Wu 1988), transformation was confirmed using reliable molecular techniques, but the patterns of DNA hybridization (see Section 11.5.2) were a little unusual and may have been misinterpreted.

11.5.1.4. Rye Floral Tiller Injection (de la Pĕna et al. 1987). In this early report of plant transformation, young floral tillers of rye were injected with DNA carrying a kanamycin resistance gene. The authors speculated that the DNA was transported through the plant's vascular system to the germ cells, where it was taken up and incorporated. They suggested that the cells that ended up forming pollen were probably transformed with this injected DNA. The end result from floral tiller injection was the production of seeds carrying a kanamycin resistance gene. Molecular analysis seemed to show the presence of an intact transgene in the rye DNA, but the most important results were only briefly described in the paper and presented as "data not shown." The term "data not shown" is used in situations where it may not be necessary to present data or images, but these data should have been presented for this work. In this paper, the authors also claim that the experiment was repeated (again, repeatability is expected for scientific reports) with similar results of recovery of transgenic rye plants. The authors wrote, "We are confidant that this simple transformation procedure can be extended to other cereals…." but this work has never even been repeated with rye. It is unclear what exactly led the authors to their conclusions, but the idea of transporting DNA through the vascular system to target the male germ cells makes one question the stability of the rye genome itself. This work was published in the journal, *Nature* which was and still is one of the world's premiere scientific journals.

11.5.1.5. Electrotransformation of Germinating Pollen Grains (Smith et al. 1994). If the ideal transformation system were available, it would be pollen transformation. What could possibly be more convenient than simply introducing DNA into pollen and then pollinating a plant to generate transgenic seed? Here is yet another report of pollen transformation that has not been pursued or repeated in over 20 years. In this report, pollen from tobacco was germinated, washed, and subjected to electroporation. Although electroporation clearly works well for transforming protoplasts and some actively growing plant tissues, it may have its limitations for stable DNA introduction into pollen. DNA in the growing pollen tube is not actively dividing and may not be receptive for foreign DNA. The authors report the optimization of DNA delivery through transient expression of gene activity, which is quite feasible. Since introduced DNA does not have to be incorporated into the

host DNA to be transcribed. Transient expression in germinating pollen is described in this paper, along with molecular analysis of some of the recovered plants. The authors reported that 40–70% of the surviving pollen (electroporation kills 35% of the pollen) displayed transient expression and one-third of the 743 recovered plants showed some activity from the transgene. This recovery rate is very high. Although proper molecular analysis of one plant appears valid, comparative analysis of more plants seems feasible and should have been presented, considering the large number of plants recovered. See Chapter 12 for methods to analyze putatively transgenic plants—we can see why these are so important in this section.

11.5.1.6. Medicago Transformation via Seedling Infiltration (Trieu et al. 2000). Although *Medicago truncatula* is a relatively unknown plant outside of the plant sciences community, it has been presented as a "model" for legumes; the plant family that includes alfalfa, peas, and all of the "beans" (soybeans, lima beans, green beans, etc.). As a legume model and potential counterpart to *Arabidopsis* (which is the unquestionable model for all plants), large amounts of resources were placed toward the development of comparable transformation technologies for *M. truncatula*. These efforts resulted in a publication describing the adaptation of the *Arabidopsis* floral dip method for this plant species (Trieu et al. 2000). Although most of the plant scientists on the planet have successfully used the *Arabidopsis* floral dip method, replication of the floral dip method described in this paper for this legume "model" have been nonexistent. Transformation efficiencies of 3–76% were reported, but it remains unclear to this day if any transgenic plants were actually recovered. The appropriate molecular analyses were set up and are accurately presented in this paper, but they were grossly misinterpreted. As opposed to the one plant analyzed from the pollen grain electrotransformation (see earlier text), many different plants were analyzed in this report. The difficulty lies in the patterns of DNA hybridization that were presented in the paper. In most cases, hybridization patterns in transgenic plants should be unique when assaying for transgene insertion; in this paper, most of the plants displayed the same single band (see the next chapter for details why this result is not expected/acceptable). The criteria to be considered in evaluating the success of transgenic plant production are not that complex. It is surprising that so many scientists are not fully aware of them.

11.5.1.7. Wheat Floral Dip (Zale et al. 2009). Wheat was targeted in this additional report using the *Arabidopsis* floral dip approach. Immature flowering wheat plants were dipped once or twice in an *Agrobacterium* suspension and allowed to produce seed. As stated earlier, this approach is so easy that, if it worked effectively with wheat or any other crop of interest, the traditional transformation approaches would be tossed to the side in an instant. In this report, two different *Agrobacterium* strains were used, which contained two different T-DNA containing vectors. The authors claimed that they recovered three different transgenic plants. Although there are no glaring errors or inconsistencies in the molecular analysis, the evidence for successful transformation is just not very clear. The autoradiographs that were presented in this work to show DNA hybridization patterns were very noisy, with spots, runs and ghost smears in the background. This background was not present in the controls, as it should have been.

Perhaps, the most surprising result presented by these authors was the recovery of the transgenic event 14C1, which resulted from a double dip. For this one event (out of the three total reported), seed set in the dipped plant was severely reduced and only one seed was recovered. Remember that no selection of any kind is typically applied to dipped plants and all recovered seeds need to either be tested or grown on selective medium. The authors reported that this single seed from this one plant was actually transformed. With the *Arabidopsis* model, which is the most efficient system known, transformation efficiencies of 1–3% are common (Clough and Bent 1998), which means that 1–3 out of 100 seed are transformed. In this report, 1 of 1 seed was transformed in wheat. In *Arabidopsis*, the floral dip technique works at low frequency, but with good consistency. Therefore, optimistic plant biotechologists reason that it should work for other plant species too. I hope that our

knowledge can be expanded and our techniques improved so that floral dip transformation frequency eventually becomes routine for other plants and that is something to strive for. Unfortunately, this seems to be another false report of floral dip transformation as the evidence is so weak.

11.5.1.8. Vacuum Infiltration of Petunia Pollen with Agrobacterium (Tjokrokusumo et al. 2000). Vacuum infiltration can be used as a method to enhance delivery of *Agrobacterium* into large leaf air spaces for agroinfiltration (see earlier text). Vacuum infiltration works when the air within the plant tissue expands under vacuum and is released as bubbles, and the former air space is then filled with the liquid when the vacuum is released. If *Agrobacterium* is present in the liquid, it will then fill the leaf air space and bind to the leaf cells for transformation. Can this approach be applied to tissues or single cells that do not posses air spaces? In this report, pollen from petunia was harvested and placed under vacuum in a suspension of *Agrobacterium*. The pollen was kept under vacuum for 20 min, the vacuum was slowly released, and the pollen was used for pollination of flowers. In a variation of this approach (not really—it was called "method 2" while pollen vacuum infiltration was called "method 1"), *Agrobacterium* was applied to the stigma surface prior to pollination with untreated pollen. For both of these unlikely approaches, the reported transformation frequency of progeny seedlings was extremely high at approximately 70%. Similar to the other reports in this section, the molecular analysis showed blurred banding patterns with very high background. The authors try to explain why the same bands were present in some transformants (they should be different), but they never try to explain how this approach could ever work, considering that pollen grains do not contain large air spaces.

11.5.2. Criteria to Consider in Judging Novel Plant Transformation Methods

As stated earlier, plant biotechnologists are constantly striving to improve existing approaches and develop new technologies that are more efficient and useful than routine methods. If new methods are to be believed to be valid, evidence should be presented that cannot be refuted; that is, data must be unequivocal. Even the most critical and ultra-conservative scientist (yes, there are a lot of those) must admit that a procedure works, if the supporting documentation and evidence is convincing. Unfortunately, the articles highlighted in the text are not well supported by the data presented and have not been replicated. These papers have received much attention and are often referenced in the scientific literature as the first or only successful report of their kinds. It stands to reason that if the new methods are effective, they would be rapidly replicated and adopted.

True success in transgene introduction in plants can be confirmed in a number of different ways. Validation of transformation is based on either the presence of foreign DNA in the plant genome and/or the expression of the transgene in the form of a new enzyme or protein. Few of these validation methods are reliable on their own; often analysis at a number of different levels is required. Below are some considerations for the main methods, which are used to confirm the transgenic nature of transgenic plants.

11.5.2.1. Selectable Marker (Resistance) Genes. One of the most common methods for false confirmation of transgene expression is to evaluate plant tissues and seedlings for resistance to herbicides (any compound that is toxic to plant tissues). Although herbicide resistance genes are almost always used as a selective agent, the levels of herbicide used for selection are often at the lower end of toxicity. This means that there is the possibility of allowing escapes, which may not contain the transgene but could still survive in the presence of the herbicide. It is rare that transformation experiments give rise to plant tissue and plants that either grow unaffected or die in the presence of the herbicide. In most cases, the recovered tissues show some yellowing or browning, indicating slight toxicity effects. The ability for plant tissue to survive in the presence of toxic agents depends on the density and vigor of the plant tissue, the medium used for growth of the target cells and the stability of the selective agent. Some selection systems—those that have been thoroughly

worked out and optimized by plant species—may be very trustworthy. But growth of tissues or seedlings on selective media is not sufficient evidence to confirm the presence and expression of an herbicide-resistant transgene.

11.5.2.2. Reporter Genes. Expression of reporter (or marker) genes results in the direct or indirect formation of a product, which can be either chemically analyzed or visually confirmed (see the previous chapter). The most common marker genes are those that can be visualized. The presence of the β-glucuronidase (GUS) enzyme encoded by the *uidA* gene is analyzed by placing the plant tissue in the presence of an artificial substrate that is broken down by the enzyme to yield a blue product. When the GUS enzyme is present, the tissues expressing the transgene will turn blue. Oftentimes, the blue product is difficult to see in green plant tissues. The chlorophyll can be removed from the tissue after treatment, for clarification. If the solution containing the artificial substrate is incorrectly modified or the plant tissue is incubated for too long, everything can turn blue, leading to false-positive results.

Another commonly used marker gene encodes the GFP, which emits a fluorescent green light if the tissue expressing the gene is illuminated with UV or high intensity blue light (Figs. 11.5, 11.8, 11.9, and 11.11). Special instrumentation is needed to detect GFP and filter sets are required. If black lights or ultraviolet lamps are used without filter sets, detection of this fluorescent protein is sometimes difficult unless the amounts of GFP protein are very high (Fig. 11.5). The main problem in detecting GFP in plants is the presence of other plant compounds that either interfere with detection or fluoresce themselves. For example, chlorophyll gives background fluorescence appearing bright red under UV or blue light. Waxes, materials in leaf hairs/trichomes, and even dirt on the leaves can fluoresce in a similar way to GFP and some filter sets can make everything look like GFP expression. The presence of the appropriate color for these marker genes must be carefully evaluated and then compared with an expected pattern for gene expression for the most accurate results.

11.5.2.3. Transgene DNA. Ultimately, the transgenic nature of a plant relies on the detection of the new transgene through DNA analysis. In some cases, DNA analysis has become so sensitive that small amounts of contaminants in the laboratory can yield false-positive results. Use of the polymerase chain reaction (PCR) must be cautiously weighed as false positives can be common with this method. In addition, PCR does not test for the integration of the transgenic DNA, only its presence in the sample. So, if there is some DNA on the leaves from an adjacent plant or the *Agrobacterium* remains in/on the plant, there will be a positive signal. PCR is a great screening tool in the laboratory, but PCR results should never be presented as the only proof of transformation.

The best method for molecular analysis of integrated transgenic DNA is Southern blot analysis (see Chapter 12 for details). Many publications present Southern blots showing the same-sized band for all clones. If enzymes are used that cut a fragment out of the transgene, a single band will be generated. A single band will also be generated if the starting DNA is from a bacteria or DNA that is contaminating the sample. If a restriction enzyme is used that cuts the foreign DNA at only one location, it will also cut somewhere in the plant DNA, producing different sized fragments from each different transformation event. More bands are typically generated from plants obtained using direct DNA introduction while *Agrobacterium*-mediated transformation yields fewer and less complex banding patterns. Regardless of the method for DNA introduction, the presence of unique band sizes and band numbers should be used to confirm transgene integration resulting from each different transformation event.

It is also important to analyze the progeny of putatively transgenic plants (see Chapter 12). A transgenic plant should pass the transgene on to progeny with Mendelian-expected frequencies. Non-Mendelian inheritance of transgenes suggests problems at some level. Some examples of potential problems might be (a) there was contamination of DNA at some level, for example,

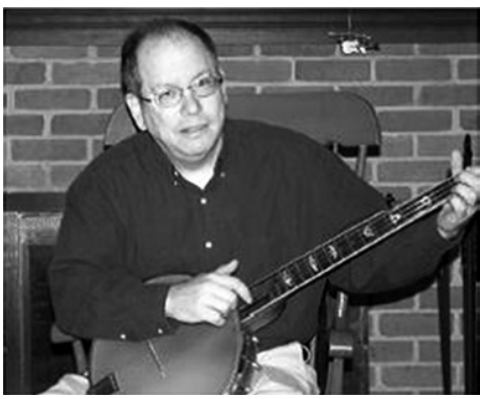
Agrobacterium carryover, (b) an endophyte of the plant might have been transformed, or (c) there was prolonged transient expression, but the transgene(s) were not integrated into the chromosome.

11.6. A LOOK TO THE FUTURE

In the early days of transgenic plant production, the major difficulty was the actual production of transgenic plants. As transformation science progressed, the procedures for gene delivery, selection, and transgenic plant production became more standardized for most plants. Transformation systems for even the most difficult-to-transform plants can now be termed "consistent but inefficient." This means that, if you know what you are doing, you can count on the production of a few transgenic plants for each experiment. Many plants that used to be difficult-to-transform are no longer even considered "difficult." So, for many plants, transformation is no longer limiting and the analysis of transgenics is the new bottleneck. Can we even analyze fewer plants if we eliminate the variation in transgene expression by developing more reliable methods to introduce the transgene into exactly the same locus (in the genome) each time? Can we follow the lead of the automotive industry by automating more of the process? Can we develop the *Arabidopsis* floral dip method for all plants? Will we solve the problems with pollen transformation? Transformation science, as with science in general, moves forward through the systematic optimization of known systems and the discovery of new approaches. Hopefully, one of the young scientists reading this chapter will take the lead to optimize or develop a new transformation technology that will eliminate one or more of the remaining bottlenecks in transgenic plant production.

LIFE BOX 11.1. TED KLEIN

Ted Klein, Senior Scientist, Pioneer Crop Genetics Research, DuPont Agriculture & Nutrition.



Ted Klein picking his banjo, claw hammer-style. Courtesy of Ted Klein.

When asked how I decided on a career in plant molecular biology, I often answer by saying that even as a student in high school, I knew that I wanted to become a soybean genetic engineer. Given that I graduated DeWitt Clinton High in the Bronx in 1972, this is a highly unlikely scenario. Of course, I am trying to make the point that it is very difficult to predict the course of one's career. I would never have predicted that I would be involved with breakthrough science that changed the course of agriculture.

Biology was my real focus in high school, and I truly enjoyed learning about the intricacies of organisms. I went on to attend McGill University in Montreal and was fortunate to major in plant science at the agriculture campus (Macdonald College). My thought was that the most important and practical aspects of biology were related to agriculture and plant development and that I would pursue a career in this area. I was drawn to learning about the interactions between organisms, especially those between plants and microbes. I found the courses in plant pathology, microbiology, and microbial ecology particularly interesting. Soil seemed to be where the real action was. I went on to do graduate work at Cornell University with

Martin Alexander, the noted soil microbiologist. My research focused on aspects of the nitrogen cycle and the organisms responsible for converting ammonium to nitrate in acid environments. As I was finishing my degree, my goal was to continue on in microbial ecology and hopefully obtain a faculty position after a postdoc. However, I had the good fortune of meeting John Sanford and learned about his concepts for genetic engineering of crops. John worked at the New York Agricultural Experimental Station in Geneva, about 50 mi from Ithaca. Driving home to Ithaca after our meeting, I was convinced that he was on to something totally new and extremely exciting.

For the next 3 years, I worked with John on the development and implementation of the gene gun for DNA delivery to cells and tissues. Our process evolved from using a real gun (air pistols and rifles) to a specially designed apparatus fabricated at Cornell's Submicron Facility with Nelson Allen and Ed Wolf. We tried to deliver small tungsten particles into anything that wouldn't move (onions, paramecia, Drosophila eggs). This was before simple reporter genes (i.e., GUS) with strong plant promoters were available. Eventually with the help of Ray Wu, we were able to bombard onion cells and show that genes could be delivered and expressed. At that time, the goal of a number of labs was to introduce genes into important crop species such as corn, rice, and soybean. We went on to collaborate with scientists at Pioneer to show that maize cells could be transformed. After working with John, I decided to do additional postdoctoral work at the Plant Gene Expression Center in Albany, California, with Mike Fromm. These were exciting times with the gene gun being applied to a number of important biological questions. We were able to directly deliver DNA into intact tissues to study transcription factors, phytochrome regulation of gene expression, and tissue specific expression. We were also able to stably transform maize, an important breakthrough for agriculture.

The gene gun is now an accepted tool in biological research with many applications in animal cell biology. Virtually, all of the transgenic corn and soybean grown by farmers was engineered with the gene gun. So as should be apparent, it is very difficult to predict the course of one's career.

LIFE BOX 11.2. JOHN FINER

John Finer, Professor, OARDC/The Ohio State University.



John Finer. Courtesy of John Finer.

I am like most of you—just an average person, trying to find my way. I was born and raised in Cincinnati, Ohio, and went to a public Cincinnati high school. In college, to please my dad, I started out as a zoology major because he thought that I should be a doctor. After my lab mate refused to help dissect a fetal pig, a standard exercise for zoology majors, I changed my major to botany because I liked growing plants. The pig dissection was awful, and I did not want to think about cats or dogs or whatever else was next. College was great fun and a valuable experience, but it was more fun than anything else. I survived the chemistry courses and actually enjoyed some of the botany courses. But, I joined a fraternity,

became involved in running the house, went to a lot of parties, and met a smart beautiful girl, who later became my wife. My grades were not very good, but they went up a lot during my senior year from the large number of study dates that I had with my future wife. I was planning on doing something in landscape nursery design or marketing, but the federal government intervened during the last semester of my senior year. The post office (an agency of the federal government) lost my tuition check and I was dropped from the auto mechanics course that I needed for graduation. When the drop was caught, I tried to re-register and was waitlisted 5 deep. This was a very difficult course to add, and there was no way that I would get in. I had just bought a piece-of-junk, rusty, barely functional, 1968 Mustang the summer before my senior year, and I really wanted to get into this class. Because it was late, I had to sign up for research hours to get the credit hours that I needed for graduation. I approached the professor of one of my favorite classes in plant pathology to ask about research hours. She said that she did not have anything and rejected me. Instead, she introduced me to a graduate student in another lab, who had a lot of projects and was looking for help. The graduate student gave me a stack of papers, which I took on a study date and read. The graduate student could not believe that I had read all of the papers and was asking him questions the next day about the work. That last semester in the lab became more fun than the fraternity parties.

The graduate student, who I worked for, saw something in me and suggested that I try graduate school since that was what he was doing. Based on his recommendation, I applied to Purdue University and Texas A&M University. I had decent Graduate Record Exam test scores, but my grades were not very good. I did not get into Purdue, but Texas A&M accepted me, on academic probation. My advisor told me the story about how the department chair came into her office, and asked her about admitting me. She shrugged her shoulders and said that she did not know what do to. After some small talk, he said to her, "let's give the poor guy a chance" and I was in! I packed my barely running Mustang with all of my stuff and drove from Ohio to Texas, sleeping at a rest stop overnight. Graduate school was very different from my undergraduate program. The classes seemed so easy and relevant to the work that I was doing. I was in the lab all the time. But, I had a lot of failures. To this day, I believe that if everything worked all the time, it would not be very much fun. The successes that I had in the lab (and outside of the lab too) were greatest if I had experienced some failures along the way. The reward was figuring things out, which I loved to do and am pretty good at.

Throughout my professional career, I have tried to be observant, find opportunities, and never overthink things. Some of my most impactful discoveries started out as a casual suggestion or an observation of structures that few others would notice. Many of my more successful collaborative projects resulted from unplanned meetings between certain colleagues and me. In science and in life, sometimes things just happen for no reason and without a plan. I find that the secret to finding success is to be observant, notice the potential, and take advantage of the incredible opportunities that are sitting in front of all of us. Because of my position within the university and success in plant biotechnology, I have had the opportunity to meet and work with some really good people. Norman Borlaug autographed a baseball for me, because he felt that one of his greatest accomplishments was helping to bring Little League baseball to Mexico (the Nobel Prize was noteworthy too). Ingo Potrykus accepted my invitation to visit my university because he is an avid birder and I set up a trip based on the migration patterns of warblers, which came through my state at a certain time of the year. I did my postdoc with Mary-Dell Chilton, who is an amazingly creative and hardworking individual. I have met governors, other Nobel laureates, presidents and a lot of incredible scientists. Most of these successful people work very hard and are totally committed to their trade, whatever it is. So, I hope that you can find your opportunities, and your passion in life and commit to it. I also hope that one of you can relate some of your past or current situation to my story and join me in the pursuit of scientific discovery.

LIFE BOX 11.3. KAN WANG

Kan Wang, Professor, Department of Agronomy; Director, Center for Plant Transformation, Plant Sciences Institute, Iowa State University.



Kan Wang. Courtesy of Kan Wang.

In the spring of 1982, I graduated from Fudan University in Shanghai, China, with a bachelor of science degree in biochemistry. Not knowing what to do next, I decided to stay in school for a few more years. I passed entrance exams and was admitted to the graduate school of Fudan University in the fall of the same year. A few weeks later, I was informed that I was selected to pursue graduate study in plant science in Belgium. While it was widely considered a privilege to be sent abroad to study, I was not at all thrilled to be going to Belgium to study plant sciences. Why Belgium? I dreamt about England, Germany, or France. My only knowledge about Belgium was that famous inspector Hercule Poirot from Agatha Christie's novel. I did not even know what the national language was in that country. In addition, I had always been fascinated by human genetics. I did not want to get into the plant area. In my mind, plant meant crop, crop meant agriculture, and jobs in agriculture meant living in a rural area, which was not what I wanted for the future of my life.

So I went to my advisor, Professor C. C. Tan, the founding father of modern genetics in China, and told him that I was not happy with the decision. Smiling, he replied, "trust me, this is the best lab in plant molecular biology."

The response from the department administration was cold, "go to Belgium to study plants, or you stay." Take it or leave it. At that time, one would do anything to go abroad. "You are so lucky to be picked, so stop complaining and start packing," said my family and friends.

Busy learning English and preparing for the long journey abroad, I was totally unprepared for what I would be doing in the new lab. The day before I left Shanghai, one of my colleagues in the plant science group passed me two review articles written by Jeff Schell (Schell et al. 1979). He said, "I heard that you are going to Belgium, this is the only paper I could find. They are studying crown gall."

The lab at Gent University was co-run by Professors Marc Van Montagu and Jeff Schell. Jeff spent 90% of his time in Koln, Germany, as the director of the Max-Planck Institute, while Marc was the major professor for most students in Gent. Marc first appeared in the lab 2 weeks after I had arrived. I was fully expecting to get an assignment from my supervisor; instead, Marc asked me what I would like to do for research. Realizing that I did not even have a clue about what was going on in the lab, he asked if I had any questions for him as he was ready to finish our conversation. I remembered that I was feeling embarrassed about being unprepared and incompetent. Just before he stepped out of the lab, I asked, "What is the function of these (T-DNA) borders?" Marc stopped and turned back to me, said excitedly, "Good question, why don't you find out?" This meeting became the first turning point of my scientific career.

Marc has the remarkable ability to attract good scientists from all over the world. In the early- to mid-1980s, the lab was enjoying success in the study of *A. tumefaciens* and the T-DNA transfer mechanism. In my opinion, the success was largely attributed to the vibrant international environment created by Marc. I, too, enjoyed tremendously and adapted quickly to such a lively culture. This experience taught me how to enjoy the diversity and appreciate the difference in science as well as in life.

My first project was designed by my research mentor Patricia Zambryski, who was a senior

scientist in Marc's lab at the time. When you were a newbie, you often got secondary projects. Like one senior student put it, "these experiments are important, but not that important. Otherwise we would be doing them ourselves."

At the beginning of the 1980s, it was known that none of the internal portion of the T-DNA is required for transfer and integration; only one of the DNA regions near the ends of the transferred DNA is required. In addition, *Agrobacterium* carrying extensive (18 kb) deletion of the T-DNA right border region was virtually avirulent on many plant species, while the deletion of the left border region does not affect the tumor-forming ability of the mutated plasmid. I was given the tasks of determining which DNA sequences around the border region are required for the T-DNA transfer. With the help of Patti and Luis Herrera-Estrella, I was able to narrow down the right border region to about 1–2 kb around the 25 bp T-DNA border sequence by DNA deletion and complementation. I was also able to show that the right border of the T-DNA functions directionally, that is, reintroducing the border sequence to a border-lacking mutant strain in the original orientation could restore the T-DNA transfer fully, but in an opposite orientation would not. The most fun part of this work was the attempt to find out whether the 25 bp conserved terminus sequences is responsible for the T-DNA transfer. This was a simple question, but technically challenging. One day, Applied Biosystems (now part of Thermo Fisher Scientific) came to the lab to sell their second commercial instrument, the model 380A DNA Synthesizer, a machine for making oligonucleotides. They promised to give a pair of oligo for free, if we purchased the equipment. I jumped to the opportunity and handed to the salesman the hand-written 25 bp sequence on a piece of paper. After 1 month, the newly purchased equipment came with two tubes of oligo labeled with my name. Cloning of the 25-bp sequence (without restriction enzyme adaptors) into a vector was a big ordeal 30 years ago, but I managed to introduce the sequence using blunt-end ligation and confirmed by using the traditional Maxam & Gilbert DNA sequencing all by myself (there was no DNA sequencing facility or services available then). As predicted, the 25-bp sequence itself can restore the T-transfer. When the work was published in Cell in 1984, I did not totally appreciate its significance, as it was given to me as an "important, but not that important" project.

My 4-year PhD study was productive, as it resulted in four research papers. Among three first-authored papers, one published in *Cell* (Wang et al. 1984), and another one published in *Science* (Wang et al. 1987). I also co-authored a *Nature* paper. Interestingly, when I graduated with the PhD degree, I still was not certain what I wanted to do with my life. I was as confused as before I started only with some gene cloning skills and enhanced problem-solving ability. After 2 more years as a postdoc in Marc's lab, I decided to move to an industry lab. I was eager to see how we could put the newly found knowledge to practice in agricultural biotechnology.

Research in early biotech industry was almost like academic research but with no limitation of funds. I also enjoyed working with a team of colleagues who shared common goals. However, after a few years I found that it was hard for me to work with enthusiasm when the projects became more narrowly defined to meet the company's short-term goals, and diverged from my research interest or curiosity. As the opportunity arisen, I moved to Iowa State University. But the industry experience provided me with invaluable experience in technology development in my effort of improving crop genetic transformation.

Since the discovery of Ti plasmid in 1970s and the understanding of the T-DNA transfer mechanism in the early 1980s, *A. tumefaciens* has become a powerful tool that revolutionized the entire field of plant science and changed the landscape of modern agriculture. Three scientists, Marc Van Montagu, Mary-Dell Chilton, and Robert Fraley, who pioneered the study of *Agrobacterium* and used it as a tool for the first genetically engineered plants 30 years ago, were the laureates for the 2013 World Food Prize. To me, *Agrobacterium* is the organism that led me into my career. This unique soil bacterium has been extensively studied and used for genetically transforming almost all plants of human interest, but I have never ceased to be fascinated by its many shades of mystery.

REFERENCES

Altpeter F, Baisakh N, Beachy R, Bock R, Capell T, Christou P, Daniell H, Datta K, Datta S, Dix PJ, Fauquet C, Huang N, Kohli A, Mooibroek H, Nicholson L, Nguyen TT, Nugent G, Raemakers K, Romano A, Somers DA, Stoger E, Taylor N, Visser R (2005): Particle bombardment and the genetic enhancement of crops: myths and realities. *Mol Biol* **5**:305–327.

Agrawal PK, Kohli A, Twyman RM, Christou P (2005): Transformation of plants with multiple cassettes generates simple transgene integration patterns and high expression levels. *Mol Breed* **16**:247–260.

Badr YA, Kereim MA, Yehia MA, Fouad OO, Bahieldin A (2005): Production of fertile transgenic wheat plants by laser micropuncture. *Photochem Photobiol Sci* **4**:803–807.

Benzle KA, Finer KR, Marty DM, McHale LK, Goodner BW, Taylor CG, Finer JJ (2015) Isolation and characterization of novel *Agrobacterium* strains for soybean and sunflower transformation. *Plant Cell Tiss Org Cult* **121**:71–81.

Chen L, Marmey P, Taylor NJ, Brizard JP, Espinoza C, D'Cruz P, Huet H, Zhang S, de Kochko A, Beachy RN, Fauquet CM (1998): Expression and inheritance of multiple transgenes in rice plants. *Nat Biotechnol* **16**:1060–1064.

Clough SJ, Bent AF (1998): Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**:735–743.

de la Pĕna A, Lörz H, Schell J (1987): Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* **325**:274–276.

D'Halluin K, Bonne E, Bossut M, De Beuckeleer M, Leemans J (1992): Transgenic maize plants by tissue electroporation. *Plant Cell* **12**:1495–1505.

Hadi MZ, MD McMullen, JJ Finer (1996): Transformation of 12 different plasmids into soybean via particle bombardment. *Plant Cell Rep* **15**:500–505.

Hashmi S, Ling, P, Hashmi G, Reed M, Gaugler R, Trimmer W (1995): Genetic transformation of nematodes using arrays of micromechanical piercing structures. *BioTechniques* **19**:766–770.

Gelvin S (2012): Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. *Front Plant Sci* **3**:52.

Graves AC, Goldman SL (1986): The transformation of *Zea mays* seedlings with *Agrobacterium tumefaciens*: Detection of T-DNA specific enzyme activity. *Plant Mol Biol* **7**:43–50.

Hess D (1980): Investigations on the intra- and interspecific transfer of anthocyanin genes using pollen as vectors. *J Plant Physiol* **98**:321–337.

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993): New *Agrobacterium* helper plasmids for gene transfer to plants. *Trans Res* **2**:208–218

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985): A simple and general method for transferring genes into plants. *Science* **227**:1229–1231.

Schell J, Van Montagu M, Depicker A, De Waele D, Engler G, Genetello C, Hernalsteens J-P, Holsters M, Messens E, Silva B, Van den Elsacker S, Van Larebeke N, Zaenen I (1979): Crown-gall: Bacterial plasmids as oncogenic elements for eukaryotic cells. In *Molecular Biology of Plants*, Rubenstein I, Phillips RL, Green CE, Gengenbach BG (eds). Academic Press, New York, pp 315–337.

Kaeppler HF, Gu W, Somers DA, Rines HW, Cockburn AF (1990): Silicon carbide fiber-mediated DNA delivery into plant cells. *Plant Cell Rep* **9**:415–418.

King JL, Finer JJ, McHale LK (2015): Development and optimization of agroinfiltration for soybean. *Plant Cell Rep* **34**:133–140.

Luo Z, Wu R (1988): A simple method for the transformation of rice via the pollen-tube pathway. *Plant Molec Biol Rep* **6**:165–174.

Maliga, P (2014): Chloroplast biotechnology: methods and protocols. In: *Methods in Molecular Biology, Methods and Protocols*. Humana Press, Totowa, NJ.

McKnight TE, Melechko AV, Griffin GD, Guillorn MA, Merkulov VI, Serna F, Hensley DK, Doktycz MJ, Lowndes DH, Simpson ML (2003): Intracellular integration of synthetic nanostructures with viable cells for controlled biochemical manipulation, *Nanotechnology* **14**:551–556.

Melechko AV, Merkulov VI, McKnight TE, Guillorn MA, Klein KL, Lowndes DH, Simpson ML (2005): Vertically aligned carbon nanofibers and related structures: Controlled synthesis and directed assembly. *J Appl Phys* **97**: 41301.

Pawlowski WP, Somers DA (1998): Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. *Proc Natl Acad Sci USA* **95**:12106–12110.

Sheen J (2001): Signal transduction in maize and Arabidopsis mesophyll protoplasts. *Plant Physiol* **127**:1466–1475.

Potrykus, I (1991) Gene transfer to plants: assessment of published approaches and results. *Annu Rev Plant Physiol Plant Mol Biol* **42**:205–225.

Rudder S, Doohan F, Creevey CJ, Wendt T, Mullins E (2014): Genome sequence of *Ensifer adhaerens* OV14 provides insights into its ability as a novel vector for the genetic transformation of plant genomes. *BMC Genomics* **15**:268.

Shou H, Palmer RG, Wang K (2002): Irreproducibility of the soybean pollen-tube pathway transformation procedure. *Plant Mol Biol Rep* **20**:325–334.

Smith CR, Saunders JA, VanWert S, Cheng J, Matthews BF (1994): Expression of GUS and CAT activities using electrotransformed pollen. *Plant Sci* **104**:49–58.

Taylor NJ, Fauquet CM (2002): Microprojectile bombardment as a tool in plant science and agricultural biotechnology. *DNA Cell Biol* **21**:963–977.

Tjokrokusumo D, Heinrich T, Wylie S, Potter R, McComb J (2000): Vacuum infiltration of *Petunia hybrida* pollen with *Agrobacterium tumefaciens* to achieve plant transformation. *Plant Cell Rep* **19**:792–797.

Trieu AT, Burleigh SH, Kardailsky IV, Maldonaldo-Mendoza IE, Versaw WK, Blaylock LA, Shin H, Chiou T-J, Katagi H, Dewbre GR, Weigel D, Harrison MJ (2000): The transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. *Plant J* **22**:531–541.

Vain P, McMullen MD, Finer JJ (1993): Osmotic treatment enhances particle bombardment-mediated transient and stable transformation of maize. *Plant Cell Rep* **12**:84–88.

Vaucheret H (1994): Promoter-dependent trans-inactivation in transgenic tobacco plants; kinetic aspects of gene silencing and gene reactivation. *C R Acad Sci* **317**:310–323.

Wang K, Herrera Estrella L, Van Montagu M, Zambryski P (1984): Right 25 bp terminus sequences of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* **38**:455–462.

Wang K, Stachel SE, Timmerman B, Van Montagu M, Zambryski P (1987): Site specific nick in the T-DNA border sequence as a result of *Agrobacterium vir* gene expression. *Science* **325**:587–591.

Yamashita T, Lida A, Morikawa H (1991): Evidence that more than 90% of β-glucuronidase-expressing cells after particle bombardment directly receive the foreign gene in their nucleus. *Plant Physiol* **97**:829–831.

Zale JM, Agarwal S, Loar S, Steber CM (2009): Evidence for stable transformation of wheat by floral dip in *Agrobacterium tumefaciens*. *Plant Cell Rep* **28**:903–913.

Zambryski PC (1992): Chronicles from the *Agrobacterium*-plant cell DNA transfer story. *Annu Rev Plant Physiol Mol Biol* **43**:4645–4690.

CHAPTER 12

Analysis of Transgenic Plants

C. NEAL STEWART, Jr.

Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

12.0. CHAPTER SUMMARY AND OBJECTIVES

12.0.1. Summary

After transgenic plants are produced, it is critical for researchers to now understand genetic composition, and biochemical and phenotypical characteristics of the new plants. It is important to confirm that the transgenes are indeed entirely integrated in the genome and how many copies are present. Various DNA analyses such as polymerase chain reaction (PCR), quantitative PCR, Southern (DNA) blot analysis, and DNA sequencing are used to understand transgenicity. The transgenic DNA is also expected to be inherited into progeny. It is also important to characterize gene expression. Gene expression is often estimated by transcript analysis by RT-PCR, qRT-PCR, and northern (RNA) blot analysis. Ultimately, we want to know how much intact recombinant protein is produced, which can be measured quantitatively using ELISA or semi-quantitatively with western (protein) blot analysis.

12.0.2. Discussion Questions

1. What are the three most common categories of transgenic plant analyses?
2. What is the role of selectable marker genes and selection, and reporter gene expression in analysis of transgenic plants?
3. In this context, what does the word "putative" mean when referring to transgenicity?
4. Why is Mendelian segregation in T1 progeny so important?
5. What are the most important controls needed for a Southern blot? What about for qPCR?
6. Of all the technological platforms discussed in this chapter, which one is most powerful? Why?

12.1. ESSENTIAL ELEMENTS OF TRANSGENIC PLANT ANALYSIS

As we saw at the end of Chapter 11, researchers have sometimes believed they had transgenic plants when they probably didn't. Or, they thought they had an efficient transformation system, but probably didn't. So, the most important starting point in analyzing so-called transgenic plants (often

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

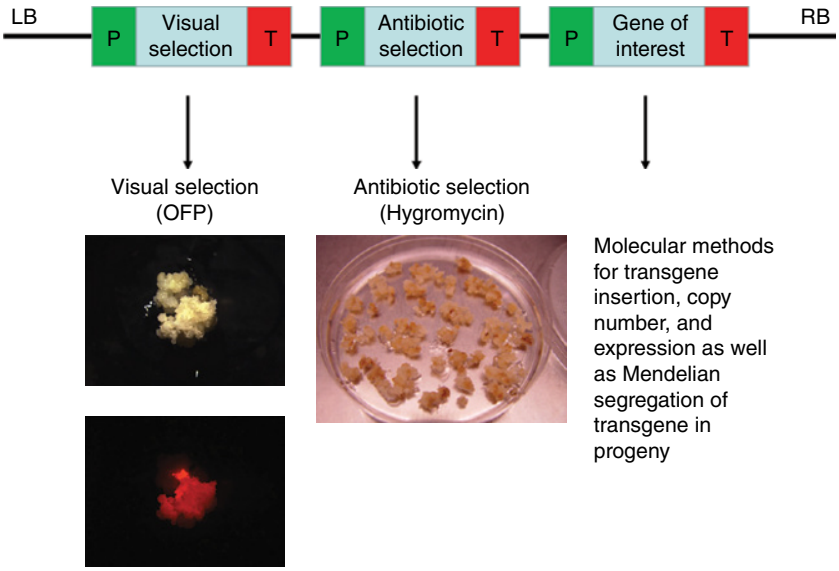


Figure 12.1. Overview of transgenic plant analysis. Several lines of evidence can be used together to assess whether the plants are truly transgenic and that the transgene of interest is expressed. Thanks to Mat Halter for assistance on this figure. (*See insert for color representation of the figure.*)

referred to as "putative" transformants) is to make sure they really are transgenic. Given enough positive evidence, we can refer to our new plants as transgenic and not putatively transgenic. As we see in Figure 12.1, there are several lines of evidence that can be used to judge whether the plants from a transformation experiment that are obtained from regenerated tissue cultures or germinated seeds from floral-dipped plants are really transgenic. Namely, if they survive antibiotic or herbicide selection, that's a good sign. If these plants clearly express the reporter gene, that's another good sign that they are transgenic. The most unequivocal evidence comes from molecular evidence—assaying for the specific DNA that is expected, and then seeing all these lines of evidence also in a fraction of the progeny plants. The "burden of proof" is typically higher in plant species or genotypes that have never or rarely been transformed or when using a novel technique.

The type of transformation vector that is typically used in research has a selectable marker gene, reporter gene, and one or more genes of interest that are targeted for overexpression or knockdown analysis using RNAi or gene silencing techniques, such as the one shown in Figure 12.1. If a company makes transgenic plants for their commercialization pipeline, they would likely not have a reporter gene. The original plants coming from a transformation experiment are termed the T0 generation. When the T0 plants are crossed or selfed, they produce T1 progeny, and so forth. In addition, researchers typically try to produce multiple independent transgenic lines (or "events"), in which the *Agrobacterium* or gene gun inserted the gene at unique chromosomal locations. The researcher can make an educated guess about which lines are independent by where they come from during the experiment, that is, if they each originated from different Petri dishes, then they are likely independent events from one another. An academic lab might be happy with a dozen or so independent transgenic events, whereas a company might want hundreds of events to find that "perfect" one for commercialization. Thus, the analysis of transgenic plants is a very big and important job. It is especially important in corporate settings because a lot of money will ride on the event selected for deregulation and commercialization.

Aside from knowing for sure that putative transgenic plants are actually transgenic and the number of gene copies, it is important to know (a) transgene expression by event, (b) if there is a correlation between transgene expression and the intended trait, and (c) if the transgenic plants are

otherwise phenotypically unchanged from the non-transgenic parent. This last issue is, again, especially important for transgenic plants to be commercialized, since farmers (and regulators) need for the plants to be substantially equivalent to currently grown crops. We will get to know some of the methods that researchers use to understand their transgenic plants at the gene (DNA) and gene expression (RNA and protein) levels. In reviewing these techniques, we will also better understand how experiments are performed and apply the scientific method to experimental setup and data interpretation.

12.2. ASSAYS FOR TRANSGENICITY, INSERT COPY NUMBER, AND SEGREGATION

12.2.1. Polymerase Chain Reaction

For some species, such as those that are very routine to transform, if plants survive antibiotic selection and express a reporter gene, then researchers might simply perform PCR on T0 events and go straight to analysis of gene expression and downstream assays. The burden of proof, based on the collective experience of plant researchers, is relatively low in these cases. Recall from Chapter 8, that PCR (Mullis 1990) utilizes thermal cycling to amplify DNA in a test tube so that enough of the target DNA can be seen in a gel assay or DNA sequenced. "Regular" PCR is performed to amplify marker gene or the gene-of-interest DNA to assess if it is really there in the sample. The great thing about PCR is that you don't need much sample—it is easy to isolate a few nanograms of DNA from many plants for PCR. The terrible thing about PCR is that you don't need much sample—contamination of the target DNA can give a false-positive result. In addition, as mentioned in the previous sample, if *Agrobacterium* is not purged from the plant of interest post co-cultivation by antibiotics, it can give a false-positive result in T0 putative transgenics. In some species and selection systems, *Agrobacterium* is completely killed; but in other systems. it can survive to haunt researchers and give them false hope.

In PCR, if the transgenic gives a band on a gel and the non-transgenic (negative) control does not show a band, then the general conclusion is that the putative transgenic is really transgenic (Fig. 12.2), but researchers generally want stronger assurance. One way to be surer is to assay for *Agrobacterium* contamination. In addition to adding PCR primers into the reaction to amplify the transgene, PCR primers can be included that would amplify an *Agrobacterium*-specific gene if present. The presence of the transgene band on the gel, and the absence of an *Agrobacterium* band, is stronger evidence for transgenicity. But, in the T0s even this PCR assay doesn't inform about transgene copy numbers.

12.2.2. Quantitative PCR

A special instrument (over and beyond the normal thermocycler needed for PCR) and fluorescent dyes that bind to DNA are added to a PCR, which amounts to quantitative PCR (qPCR, which is also known as real-time PCR). If the correct controls are used, and the sample DNA template dose is tightly controlled, qPCR can be used to give a quantitative assay about how many gene copies, relative to the control, is in the starting sample, hence the number of transgene inserts in each transgenic plant (Yuan et al. 2007). As DNA is amplified during PCR, the special instrument reads the fluorescence signal that corresponds with how much PCR product (amplified DNA) is being produced during each cycle DNA amplification via PCR (Fig. 12.3). Early cycles in PCR yield an exponential amount of amplified DNA product (doubling) after each heating and cooling (thermal) cycle. As reagents are used up, the reactions enter the linear phase, then finally the plateau phase. As more and more DNA is amplified, there is a cycle number that represents the boundary between the exponential phase and the linear phase—the threshold (Fig. 12.3). This cycle number at the threshold is called the Ct number—the exact cycle where the curve goes from exponential to linear. For each sample, there is a characteristic Ct number, and these can be compared (knowns vs. unknowns) to infer how much DNA

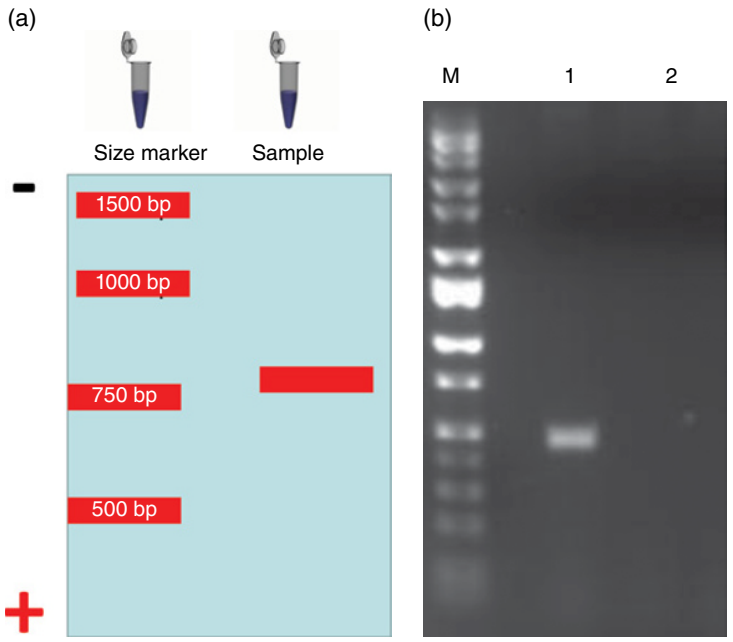


Figure 12.2. PCR analysis. (a) Graphical representation of what occurs when a DNA size marker and a PCR-positive sample is loaded into an agarose gel and subjected to electrophoresis. The larger fragments move more slowly through the electrical field and are retained toward the top of the gel as fragments are size-fractionated. (b) A photograph of a real gel containing a size marker (M) that contains DNA fragments of known size, and 2 samples. Sample 1 is PCR-positive and sample 2 is PCR-negative.

(transgene copy number) was present in the initial sample. This is a powerful method that, when used correctly and with all the controls and replicates called for, can be used to reliably estimate transgene copy number in a transgenic plant. The controls that are needed are, like regular PCR, a known non-transgenic plant, and, at least, a known positive sample that contains the equivalent of one transgene copy. Even better is also a known single-copy transgenic plant sample. This way, the Ct numbers of the unknowns can be compared to the known, single-copy standard. qPCR is a powerful method for gene copy number and transgenicity, but not the "gold standard" of DNA assays.

12.2.3. Southern (DNA) Blot Analysis

Southern blot analysis is considered to be the ultimate analysis for transgenicity and copy number. The reason why researchers might prefer qPCR over Southerns is that it is more rapid and amenable to assaying a large number of transgenic plants. qPCR also requires much less DNA per experiment. That said, for "new" plants and when the researcher must have the best data possible, there is no real substitute for Southern blot analysis.

In the 1970s, Professor Edwin Southern invented a method to transfer DNA from gels to paper (then later, nylon) membranes where the DNA could be probed with radioactive (then later, nonradioactive) labeled DNA (Southern 1975). Thus, Southern blotting shows the size of a fragment of DNA being probed, and given the correct methods and controls, the number of transgene inserts in a transgenic plant.

The method typically requires tens of micrograms of high-quality genomic DNA per sample and takes a week to perform one experiment of up to 40 samples. The DNA to be analyzed is cut in specific locations by a carefully chosen restriction endonuclease (check back to Chapter 8 for how restriction enzymes work). The cut DNA is separated on an agarose gel via electrophoresis to separate the fragments by size (Fig. 12.4). Because of the high number of fragments (thousands), the

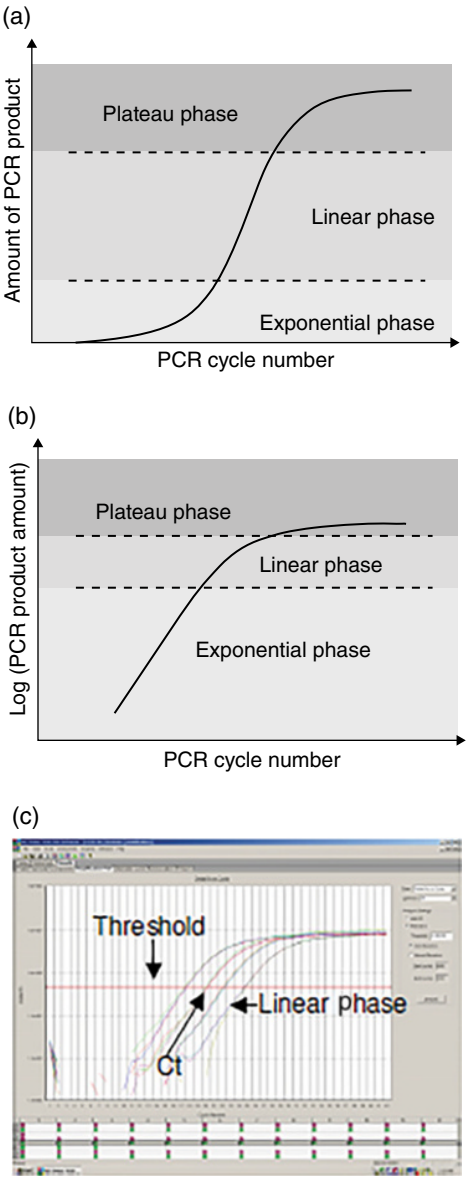


Figure 12.3. The dynamics of qPCR and analysis. (a) Theoretical plot of PCR cycle number versus PCR product showing the phases of DNA amplification. (b) Another view of the phases, but where PRC product is expressed in logarithmic terms. (c) The same scheme as panel (b), but with actual data of four samples are shown. The amount of target template decreases in the samples going from left to right as shown by respectively increasing cycle threshold (Ct) numbers. Ct is defined as the cycle at the boundary between exponential and linear phases. By knowing exactly how much DNA is in certain samples in the beginning, we can infer the amounts of DNA (and hence copy numbers) in the unknowns. This figure is reprinted with permission from Yuan et al. 2006.

digested plant genomic DNA appears as a smear (Fig. 12.4). The fragments are then transferred to a nylon membrane via capillary action so that they are in the exactly correct orientation with each other (Fig 12.5). The nylon membrane allows for a sturdy matrix for handling the DNA in the Southern blot hybridization procedure. Blotted DNA can be chemically denatured to expose each of the two DNA strands for complementary hybridization with a labeled (so we can see where it binds in the blot) DNA probe. The DNA probe is typically part of the coding region of the transgene or

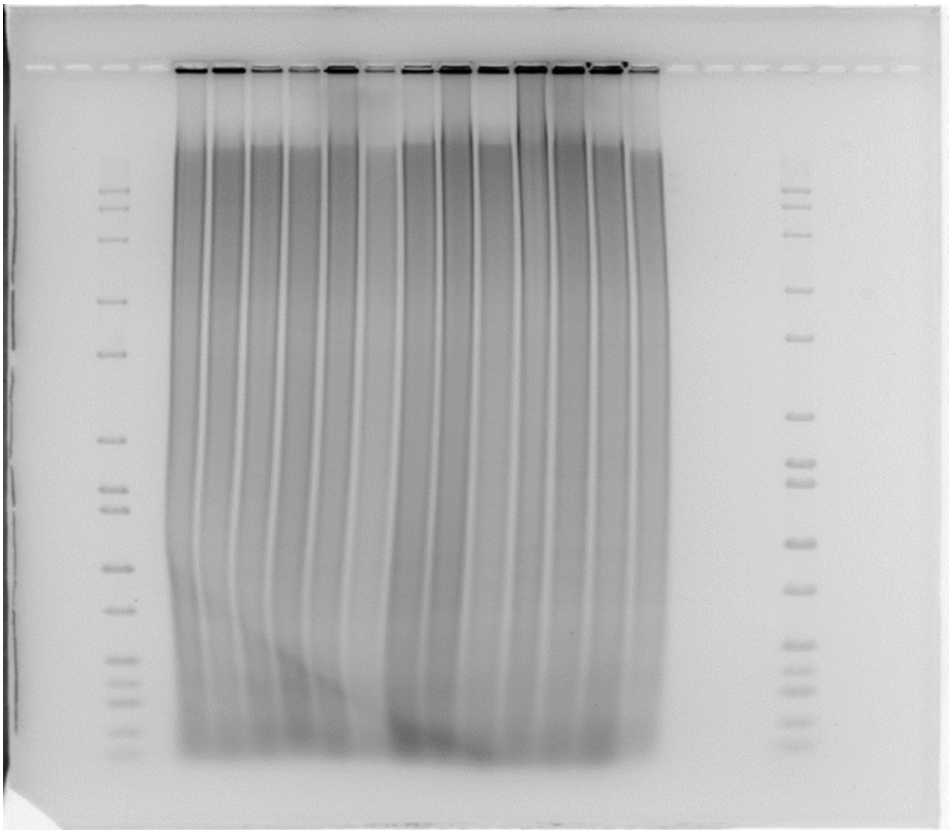


Figure 12.4. Thirteen samples of plant genomic DNA are completely digested by a restriction endonuclease and subjected to agarose gel electrophoresis to separate the DNA fragments according to size. The DNA is stained. Flanking these samples are "apparently" empty lanes and flanking these lanes are DNA size markers. One or more of the apparently empty lanes contains cut plasmid DNA that can be used as a positive control in the Southern blot analysis. The DNA will be transferred to a nylon membrane that can be probed by a labeled DNA molecule of interest.

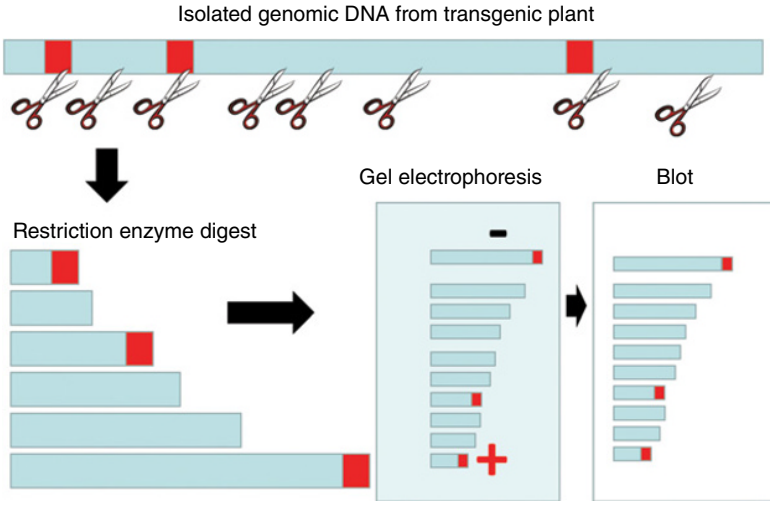


Figure 12.5. When genomic DNA is digested using a restriction endonclease, it results in a smear as shown in Figure 12.4. DNA is cut into many fragments as illustrated here. The small dark squares do not represent anything real biologically, but are used to help the reader track the fate of DNA fragments as they are digested, electrophoresed, and blotted onto a nylon membrane. Thanks to Mat Halter for assistance with this figure.

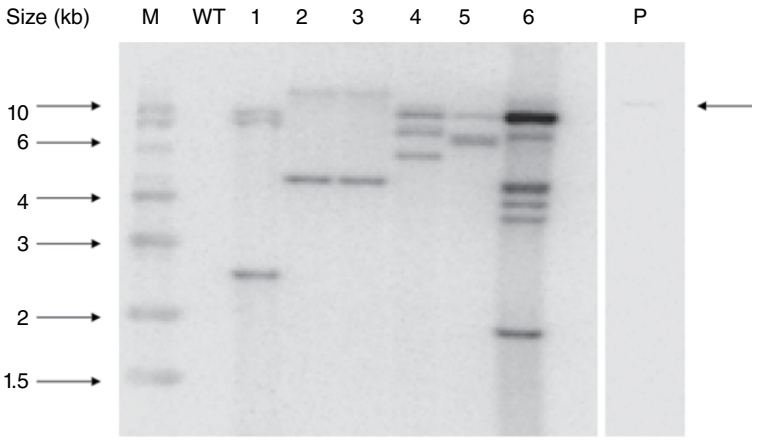


Figure 12.6. The raw data of part of a Southern blot experiment (superfluous lanes were removed for simplification). BamHI-digested genomic DNA was loaded in each of the plant lanes: WT (non-transgenic wild-type) and 1–6 (each putative independent transgenic T0 plants). M represents a DNA marker and P represents the plasmid control sample containing the gene of interest, which will bind to the DNA used as a probe; the arrow points to the faint band. It appears as if lanes 1–6 represent 5 independent transgenic plant events.

marker gene, but researchers can probe virtually any DNA. When the DNA probe finds a complementary fragment on the blotted DNA during the hybridization segment of the experiment, it shows up as a size fractionated band (Fig. 12.6).

If the goal of the experiment is to determine whether the putative transgenic plants are really transgenic for the gene of interest, and determine the numbers of loci the transgene(s) are inserted (copy number), then the probe DNA used will be from, and targeted to, the transgene. One crucial choice is which restriction endonuclease is chosen to digest the plant genomic DNA that gets separated on the gel and gets blotted to the nylon membrane. For a copy-number experiment, the restriction enzyme must cut on just one side of the probe site in the T-DNA (or plasmid, if biolistics is used), which is illustrated in Figure 12.7. In this particular case, BamHI is used as the restriction enzyme, since it will show how many copies (or inserts) of the transgene in each plant sample. We know exactly where BamHI (and other enzymes) cut in the T-DNA because it would have mapped out when the plasmid was constructed. But we don't know where BamHI sites are in the plant chromosomal location—where the T-DNA gets integrated—and that is why BamHI is diagnostic in this case. If EcoRI had been used to cut the genomic DNA samples, we can see that the same size fragment would be detected by the probe in the Southern blot (Fig. 12.7), and thus we would know nothing about copy number or the number of independent transgenic events recovered from the transformation experiment.

Interpreting Southern blot experiments can be somewhat subjective, and we will try to interpret the results shown in Figure 12.6. We see a faint band in the plasmid (positive) control lane. We should have loaded a bit more plasmid DNA in the gel, but this result is satisfactory. We also see that there is no hybridization to the non-transgenic plant (WT) sample, which is our negative control, which is also good. That means our intended positive and negative controls are doing what we expect. When we chose the six putatively transgenic T0 plants to analyze, we initially thought they would be six independent events, but we see in the blot that samples 2 and 3 have the same banding pattern. That tells us that they are most certainly duplicates of a single transgenic event—we can call this event 2–3, and the other events will correspond with the sample number. Clearly, event 1 has at least one insert represented by the 2.5 kb band. There might be two other inserts that are represented by high-molecular-weight bands. But these are faint—they could represent partially digested DNA that hybridized with the probe. Event 2–3 is probably a single insert event represented by the dark band at 5 kb. Event 4 has three T-DNA inserts and event 5 has two,

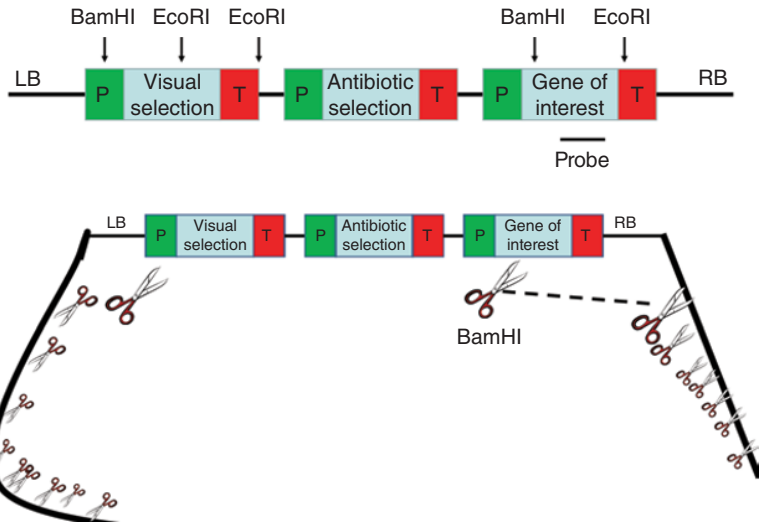


Figure 12.7. A schematic showing the T-DNA construct (top) and rationale behind the choices and setup of the experiment whose results are shown in Figure 12.7. In this vector, the BamHI and EcoRI restriction sites are shown, as well as the location of the probe DNA (top). When the T-DNA gets integrated into a plant genomic locus on a chromosome (bottom), the scissors represent actual cutting sites and some of the DNA fragments generated. Only the fragment represented by the dashed line will be hybridized by the probe in the Southern hybridization. Thanks to Mat Halter for assistance with this figure.

maybe three inserts. Note that for event 5, there might be two bands co-segregating at approximately 6 kb. Event 6 has at least six copies of the transgene, and is, by far, the "messiest" transgenic plant analyzed.

12.2.4. Segregation Analysis of Progeny

After the qPCR or Southern blot experiment is performed on T0 transgenic plants, the plants will be selfed or cross-pollinated to produce progeny T1 plants from each transgenic event of interest. Indeed, these same DNA analyses will likely be performed on selected T1 plants. However, arguably even more important is performing progeny analysis using the selectable marker- or reporter genes as tools.

The transgene, in almost all cases, is integrated in only one homologous chromosome locus. Recall, that when we think about Mendelian genetics, we know that for any gene or locus on one chromosome, there is a counterpart gene (allele) or locus on the corresponding homologous chromosome for diploid (which we assume here) plants. In "normal" genes under simple dominance, the gene state can be homozygous (dominant or recessive) or heterozygous. We can assume here that the new locus where the transgene lands also will have a counterpart on the homologous chromosome, and we will also assume the transgene does not land within another gene (although this can happen). Instead of being heterozygous, the transgene state is called "hemizygous" since there is no transgene of any type on the homologous chromosome. Therefore, in a T0 transgenic event with one transgene insert, the transgene will be hemizygous. If the plant self-fertilizes, then we can expect 1 : 2 : 1 segregation of the transgene, which will appear as 3 : 1 Mendelian segregation if the transgene is dominant (expressed), which it will be for a marker gene. For the tobacco plants in Figure 12.8a and the canola plants (Fig. 12.8b) that are T1 segregating progeny, the GFP gene appears to segregate 3 : 1. Therefore, in each of these cases, we can infer that the relevant parent (T0) events for these transgenic plants were both single copy for the T-DNA (transgene) insertion. As we will see later, we could also have used a selectable marker assay that would use its selection agent

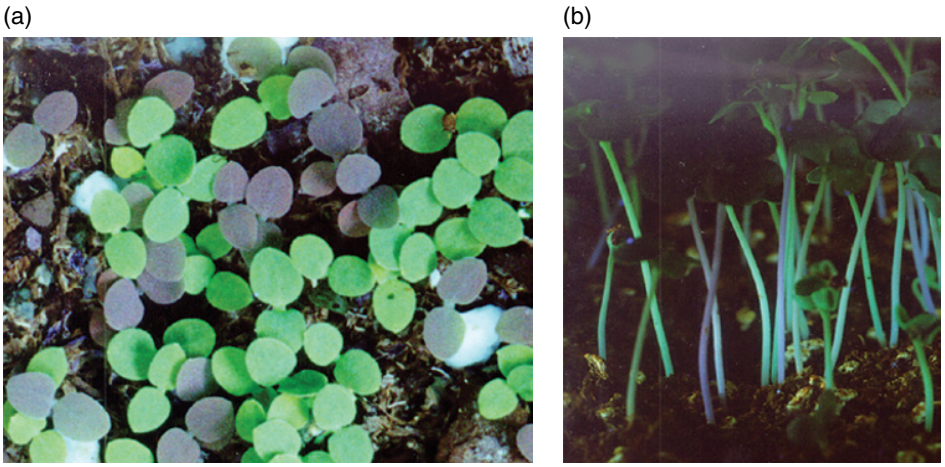


Figure 12.8. Segregation analysis of T1 transgenic (a) tobacco and (b) canola seedlings that have a single insert of a green fluorescent protein (GFP) gene. Under a UV light, the transgenic plants fluoresce green and the non-transgenic plants fluoresce red. The transgene presence and the single insert into the genome are confirmed by the Mendelian 3 : 1 segregation pattern in both of these cases. (*Source*: Reproduced with permission from Harper et al. (1999). (*See insert for color representation of the figure.*)

to help see which plants are transgenic and which ones are not. Typically, it is difficult to pick out the hemizygous plants from the homozygous positive plants in these kinds of assays. Recall from Chapter 2 that if there are two inserts (e.g., the dihybrid cross), then we would expect a 15 : 1 Mendelian segregation ratio for the marker phenotype in T1 plants. If the experiment yields more than two inserts, then a large number of T1 progeny are required for the segregation assay. The larger the number of inserts, the more difficult it is to make inferences from progeny analysis. Typically, biotechnologists are most interested in determining which T0s have single inserts, since using these plants makes the downstream analyses simpler.

12.3. TRANSGENE EXPRESSION

When assaying for transgene expression, there are two targets for molecular analyses: transcript and protein.

12.3.1. Transcript Abundance

Transcript abundance is most often used as a proxy analysis for overexpressed transgene expression. We all know that most of the time, the recombinant protein is what we are ultimately interested in assaying, but the transcript is much easier to analyze. All three of the methods to analyze the transgenic mRNA have direct comparators to the DNA analyses described earlier. Instead of doing "regular" PCR, the transcript can be analyzed by reverse transcriptase-PCR (RT-PCR). Instead of qPCR for DNA, the researcher can perform qRT-PCR. Finally instead of Southern blots, the researcher can perform northern blots (which is also a nod to Edwin Southern). In each of these assays, the researcher would isolate either total RNA or just mRNA. Next, the RNA would be converted to cDNA using reverse transcriptase—the "RT" in the terms given already, which was explained in Chapter 7.

Essentially, all the advantages and disadvantaged of the DNA analyses are relevant to their corresponding RNA analyses. Almost no one trusts RT-PCR as a stand-alone transcript assay currently given the power of qRT-PCR for a quantitative and statistically relevant assay of transgene transcript levels (Yuan et al. 2006). qRT-PCR gives relative transcript abundance data—in which an endogenous and/or transgene expression is measured against a standard "housekeeping" gene

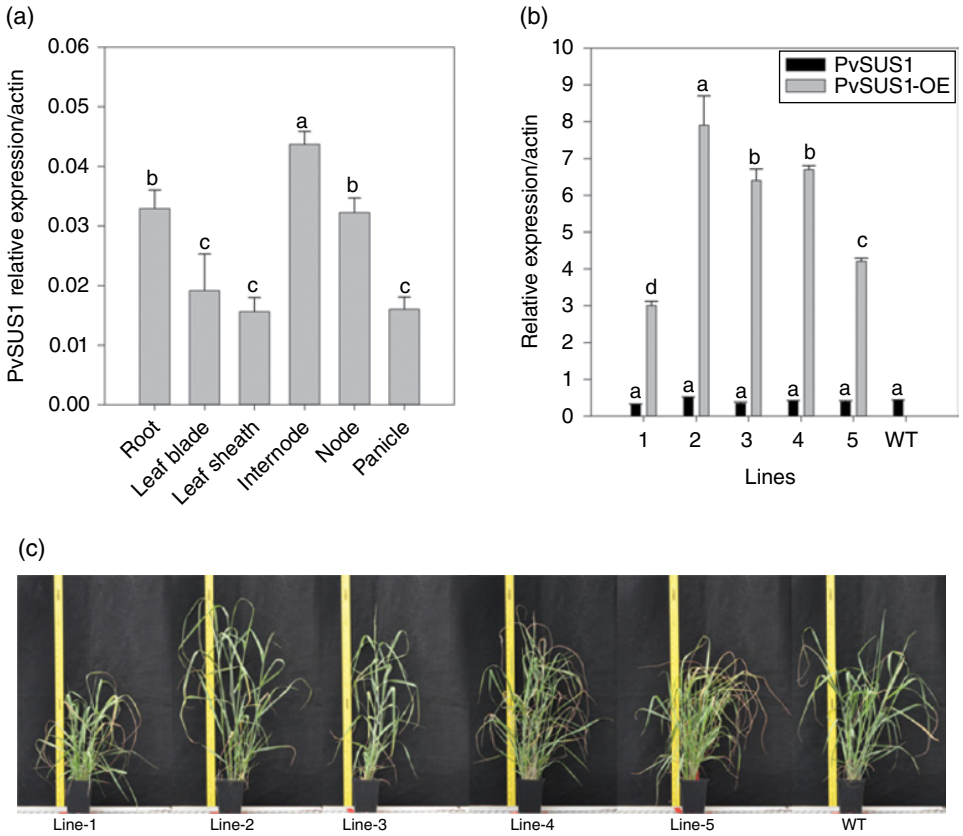


Figure 12.9. Relative transcript abundance and phenotypes of transgenic switchgrass plants overexpressing a switchgrass sucrose synthase gene. qRT-PCR analysis was performed for (a) the endogenous *PvSUS1* gene for various tissues in non-transgenic plants, along with (b) the endogenous and overexpressed *PvSUS1-OE* gene in transgenic tillers. There appears to be an association between transgene expression and (c) biomass, where the highest expressing line (2) plants were the largest, followed by lines 3 and 4. (*Source*: Reproduced with permission from Poovaiah et al. (2015).)

in each sample. Using this assay, the expression of a gene might be correlated with a trait (Fig. 12.9). qRT-PCR has easily become the most-used transgene expression assay given its accuracy and speed.

Northern blot analysis is relatively seldom used—it takes a lot of high-quality RNA to perform the assay. The cDNAs are blotted onto a membrane and a DNA probe is used that is complementary to the transgene. The difference between Southerns and northerns is that since DNA is a lot more stable than mRNA, it is harder to get sufficient quantity of mRNA than DNA. An example of a northern blot analysis is shown in Figure 12.10. Here, we see that it is important to use a negative control and an internal control—the latter is typically a transcript, a "housekeeping gene" that is stably expressed in the tissue of interest. An internal control of this type would also be used in RT-PCR and qRT-PCR, so that the results of the transcript of interest can be compared with a transcript assumed to be stably expressed in a predictable fashion.

12.3.2. Protein Abundance

If a common recombinant protein is to be analyzed, or the researchers have plenty of patience and resources, then they might choose to also analyze the recombinant protein of interest as well as the transcript. Either of the two commonly used protein analyses—western blots or enzyme-linked

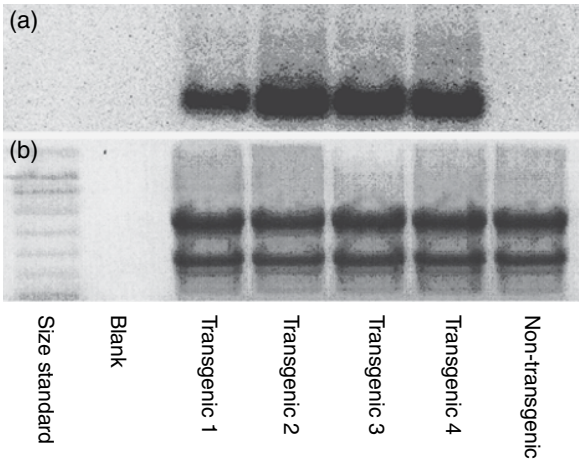


Figure 12.10. Northern blot experiment with four transgenic plant sample and a sample from a non-transgenic control. (a) Relative transcript abundance of the transgene mRNA is shows that transgenic event 1 has lower expression than that of the other three transgenic plant events. (b) Hybridization with an actin gene that is assumed (and demonstrated here) to be expressed at the same level among all plant samples, including the non-transgenic control. Here, the actin gene is used as an internal control.

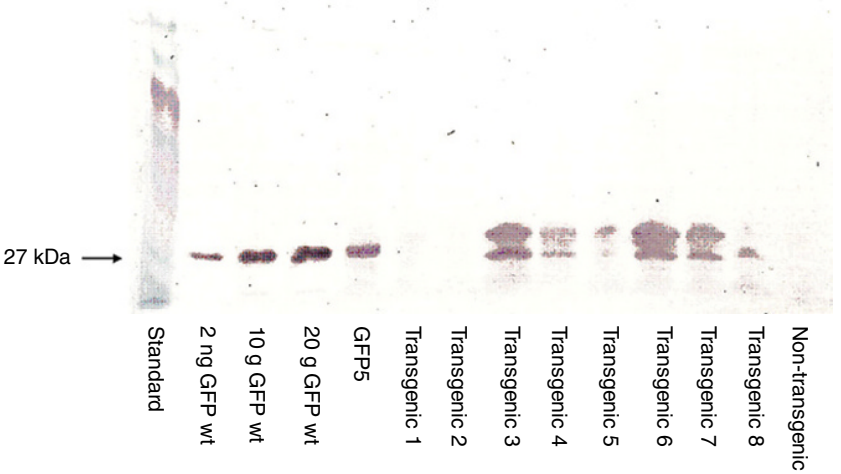


Figure 12.11. Semiquantitative western blot analysis of transgenic plants synthesizing the green fluorescent protein (GFP). From left to right, the various samples included a protein size standard, three known purified GFP quantities that form the standard curve, GFP5, which is the version expressed in the transgenic plants, and samples from eight lines of transgenic plants along with a non-transgenic plant negative control. The GFP-specific antibody was purchased commercially and visualized using an acid-phosphatase secondary antibody after staining.

immunosorbant assays (ELISA) require that protein-specific antibodies be produced, which takes time and money. Western blots, like Southerns and northerns, have the advantage of visible observation of the novel product—in this case, the recombinant protein on a blot. Typically, the protein of interest is detected using stacked antibodies and a fluorescent or biochemical tag (Fig. 12.11). Western blots are only semiquantitative, whereas ELISA are quantitative assays. Like westerns, ELISAs require specific antibodies to be produced, and they are configured on a microtiter plate—typically that has 96 wells. There are a number of commercially produced ELISAs and kits that use

the same sort of approach. Some simplified +/– kits, say for Bt proteins, are simply "dipsticks" that are mixed with a plant sample to show whether they are transgenic or not, which is very useful in certain situations.

12.4. KNOCKDOWN OR KNOCKOUT ANALYSIS RATHER THAN OVEREXPRESSION ANALYSIS

In all instances described earlier, transgene overexpression is discussed. The transgene is typically overexpressed because it is placed under the control of a strong constitutive promoter. In some instances, researchers might wish to use tissue-specific or inducible promoters that might yield a more subtle transgene expression pattern. Sometimes, however, we want to knockdown the expression of a native (or endogenous) protein or even knock out the gene entirely.

Knockout analysis has gotten to be simpler with the advent of CRISPR (discussed in Chapter 17), wherein an endogenous gene in a plant can be made to lose its function.

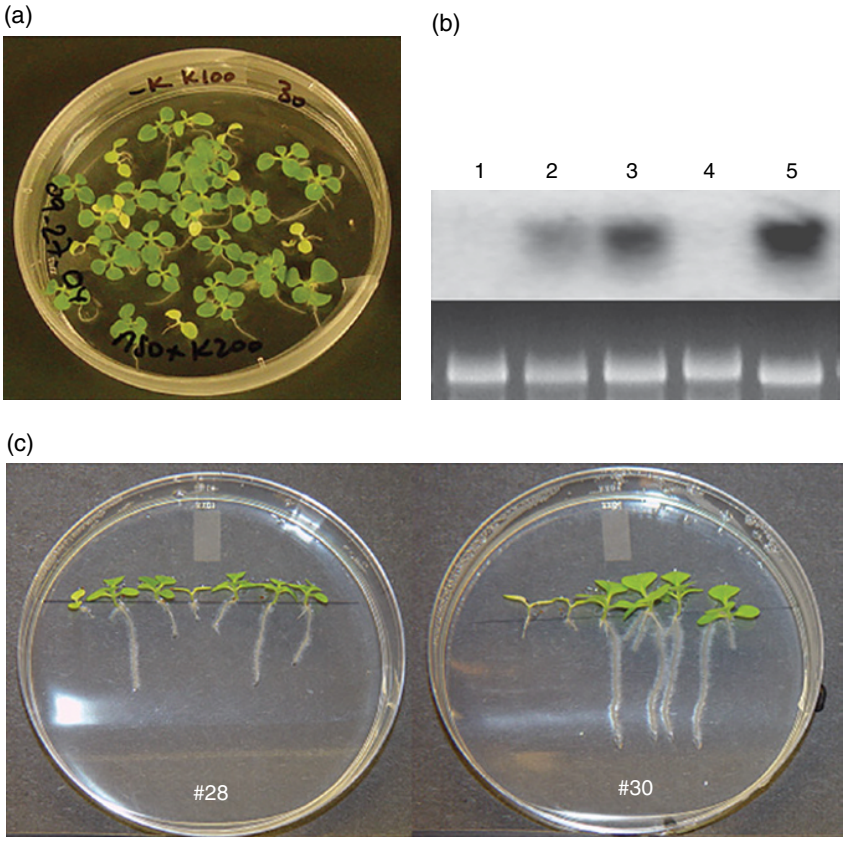


Figure 12.12. Multiple analyses of kanamycin-resistant T1 transgenic arabidopsis plants. (a) Segregation analysis of transgenic event 30 in a Petri dish containing 200 mg/L kanamycin. The large plants are transgenic and the small, pale plants are non-transgenic null segregants. (b) Northern blot analysis using the gene of interest as a probe to analyze transcript of (1) a non-transgenic control plant, and four transgenic events using a ribosomal RNA as an internal standard. (c) Kanamycin-resistance assay of two high-expressing transgenic events (28 and 30) under 200 mg/L kanamycin. Here, the T1 progeny are segregating, and this Petri dish assay allows quantitative data to be taken for root growth. The very small plants are non-transgenic null segregants. (*Source*: Reproduced with permission from Mentewab and Stewart (2005).)

In the case where gene knockdowns are desired, the transgenic plants are produced with the gene of interest targeted by RNAi (see Chapter 8), but a selectable marker gene and/or reporter gene would still be overexpressed. In either knockouts or knockdowns, the same sorts of DNA, RNA, and protein analyses can be performed. The only difference is in expression analyses, where we look for changes in the expression of endogenous genes, which we hope are lower (using RNAi) or absent (using CRISPR).

12.5. THE RELATIONSHIP BETWEEN MOLECULAR ANALYSES AND PHENOTYPE

We end the chapter with a "real-life" example wherein real data are presented to analyze. In Figure 12.12, the results of an experiment performed by Mentewab Ayalew (Mentewab and Stewart 2005) are presented. Here we sought to understand the effects of a potentially novel antibiotic-resistance gene from plants. The gene (an ABC transporter) was overexpressed in multiple transgenic tobacco lines. It became clear from Southern blot data (not shown) that several single-insert lines were interesting for analysis—especially events 28 and 30—which was confirmed by progeny analysis on kanamycin-containing plates (Fig. 12.12a). Some of these events seemed to give strong kanamycin resistance and were analyzed further. The reader can see that there is an apparent association between transgene expression (the northern blot results in Figure 12.12b) and their survival in kanamycin-containing media. In performing such "expression vs. trait" experiments, evidence for gene function can be gathered. Transgenic plants are especially valuable as tools in functional genomics.

LIFE BOX 12.1. HONG S. MOON

Hong S. Moon, Research Scientist, Monsanto Company.



Hong S. Moon. Courtesy of Hong S. Moon.

A plant biotechnology course offered in college in South Korea where I grew up opened my eyes and led me to a new field of science exploring the inside of plants. Furthermore, the course even offered insight toward manipulating plants for improvement. I was highly inspired by this course and wanted to take a deep dive into plant biotechnology. After my several applications for a graduate program, I fortunately got into a master's program studying plant biotechnology at the University of Tennessee. I came to the United States with a big dream of being a genius plant biotechnologist. My master's program was arduous as there were some steep hills to climb. Speaking a different language other than my mother language all day long was challenging. One of my friends recalled that he thought I knew only two English words: food and lunch. Indeed, at the beginning of my time in the United States, he observed that I reacted only to these two words. Studying plant biotechnology was also challenging as I had to learn a lot of molecular biology skills in the laboratory. Despite all the

challenges, I was very fortunate to have several good colleagues who were willing to help me in all aspects of research and life, and a very supportive advisor, Neal Stewart, who shepherded me with great patience through my Ph.D. program and helped me to become an independent researcher.

My master's thesis was about the consequences of gene flow from genetically modified canola to conventional canola. When I transformed and characterized plants, I don't remember how many days and months I had spent to just get a clean gel picture. But it was a great opportunity to learn about environmental risk assessment with genetically modified plants through this project. After finishing my master's program, I had an opportunity to work on developing biological containment systems to minimize unwanted pollen-mediated gene flow during my Ph.D. program. As typical in Ph.D. programs, I had to pass a preliminary examination in order to become a Ph.D. candidate. I was focusing on only my research area but nothing else. As a result of this biased preparation, I failed on the examination due to lack of my basic but essential knowledge in plant biology and biotechnology. When I was depressed, my advisor shared his own story when he was a graduate student, and that his first preliminary examination was not successful. He told me that he passed the exam on the second try. His story encouraged me to overcome that tough moment. Although it was not a pleasant moment at the time, this experience was tremendously helpful for me to understand more about plants and plant biotechnology.

During my graduate studies, I became a true believer that plant biotechnology can hugely benefit our lives. With all these experiences, I became more passionate about plant biotechnology. Considering the rapid growth in the global population, I believe that plant biotechnology plays a critical role in producing more food. After my first exposure to plant biotechnology in college, I have been very blessed to have the opportunity to work on improving plants every day. If I were ever to go back to graduate school, I would choose multidisciplinary training in plant sciences, including agronomy, weed science, soil science, and plant physiology, although my focus would still be plant biotechnology. I think plant biotechnology is not just a part of plant science, but it is an overarching theme for the future.

LIFE BOX 12.2. NEAL STEWART

C. Neal Stewart, Jr. Professor and Racheff Chair of Excellence in Plant Molecular Genetics, University of Tennessee.



Neal Stewart accompanied by a 1974 Gibson Hummingbird in the "Sound Dungeon" recording studio. Courtesy of Neal Stewart.

My early childhood years in the 1960s were spent on a small family farm not far from the proverbial Mayberry in North Carolina. My grandfather was the farmer and all of his daughters, including my mother, built houses on adjoining property, like satellites around the home planet. As suburbia encroached and grandpa grew ill and died, the 1970s rolled in and life went on. Farm life and nature were chief interests in my childhood and formative years, but also were hot rods. Biotechnology is kind of like that too—a combination of nature and technology that is somewhat of a paradox.

In college, I majored in horticulture and agricultural education. I figured I was either

going to grow flowers or teach. In those days, I had (mistakenly) told myself that I was not smart enough or a good enough student to really go into science. A spiritual awakening at the end of my college years coupled with a few-year stint of teaching middle school (in-school suspension of all things!) convinced me that science might be out of reach after all. During this time in the 1980s, I was also a singer-songwriter—an indie artist wannabe. Amongst all these seeming failures, my fairly recent (and pregnant) wife and I decided to pack our bags and head off to graduate school where I was fortunate enough to work with ecophysiologist Erik Nilsen at Virginia Tech for masters and PhD degrees. I still wonder why he took me under his wing—I was a babe in biological research, with no experience in science. His nurturing and the support of my wife got me through the MS in ecology. With a bit more confidence, I decided to add DNA into the mix of ecology and studied the population genetics and phenotypic plasticity of cranberry. I can still recall the laughter among my peers when I said in the early 1990s that I wanted to be a molecular ecologist. No one there had ever heard of such a thing, but my choice of phraseology was validated when the journal *Molecular Ecology* was begun.

Severely bitten by the DNA bug by this time, I was also fortunate to gain entrance into Wayne Parrott's lab at the University of Georgia. I think he was having a hard time finding a well-qualified postdoc, and I was foolish enough to naively launch into the soybean transformation project he had going. Soybean transformation was notoriously difficult, and I had absolutely no experience in transgenics. But again, I was fortunate to team up with Donna Tucker in Wayne's lab. She is one of those few gifted scientists who has the "golden hands" in the craft of tissue culture and an eye to select the right stuff and throw away the wrong stuff.

Biosafety research was then a natural area for me to combine transgenics and ecology—something I began in Wayne's lab and have continued on during my career as a faculty member. Initially, I was a GMO skeptic—I was convinced that there would be ecological downsides of releasing trillions of transgenic plants into the environment. But by 2004 when I had written *Genetically Modified Planet*, I had become convinced by reams of data that there were far more current and potential environmental benefits from biotechnology than risks. I still loved nature and could clearly see how the technology could make farming more environmentally-friendly.

My lab is now full of exciting young scientists who are doing all sorts of projects—from bioenergy to environmental phytosensing to weedy plants genomics to synthetic biology. My work-related happiness is mainly in doing my part to make their dreams come true.

At various times, I still worried that I was "not smart enough" for science. I worried that research funding would dry up and worried about lab personnel—finding them and then keeping them funded. Worry is a waste of time. All the stuff I worried about was moot. My advice to young scientists and students of all types is this: follow your dreams. Find your focus and the right people to help make your dreams come true. The other thing I would say to young scientists is not to become so focused on the lab that you forget to live. One thing I decided to do relatively recently—after a 20-some-year hiatus—is to reboot my songwriting "career." Given that Nashville is just a relatively short jaunt from my home, my personal country music renaissance has been birthed (you can find my music demos on reverbnation.com that were mostly recorded in my basement studio: the Sound Dungeon, or some acoustic tracks on youtube.com). I've found that this musical pursuit has exercised a different part of my brain (and being) than science, which I think is healthy. I actually think that my scientific thinking and writing has improved from my musical thinking and writing. I've also enjoyed some modest success in this venue, which makes me feel happy. And, by all means, in addition to being a productive scientist, there is nothing wrong with being happy.

LIFE BOX 12.3. NANCY A. REICHERT

Nancy A. Reichert, Professor, Department of Biological Sciences, Mississippi State University.



Nancy A. Reichert. Courtesy of Nancy A. Reichert.

In my high school valedictorian speech, I quoted Robert Frost's poem "The Road Not Taken." This poem was certainly not unique to a valedictorian speech and probably quite cliché; but looking back on my career, I realized that I had actually chosen a road and that my route was heavily influenced by people I encountered along the way.

For me, closing routes assisted in choosing my road, albeit rocky at times, through science. I attended the University of Wisconsin-Madison (UW) and quickly realized that I probably shouldn't major in music (voice) because of my stage fright, or in nursing (although there was a strong tendency for females in my family to be nurses) because I was afraid of needles. When I took Ken Todar's "Elementary Microbiology" course at UW in 1976, I felt like I found my niche and planned to work in a food quality control lab, postgraduation. Since those jobs were less abundant in Madison, the city I wanted to stay in, I sought other jobs in science—which was pivotal to my future direction. I was first exposed to plant tissue culture when Leigh Towill (USDA) took a chance and hired me as his research technician (UW Department of Horticulture) where I worked on potato meristem-tip culture for virus elimination, and cryopreservation. He collaborated with John Helgeson (USDA; trained under Folke Skoog) who was developing potato protoplast isolation, culture, and fusion procedures; learning under his technician G.T. ("Gerry") Haberlach, I was allowed to assist in the development of these procedures, which made me realize how much I enjoyed plant tissue culture.

A few years later, Helgeson's UW colleague, John Kemp, hired me as his technician at a new biotechnology company in Madison—Agrigenetics Advanced Research Lab. Being one of the pioneer biotechnology companies at the time, there was a race to see who could generate the first transgenic plants, so the excitement was palpable. Working under Kemp's direction and with guidance from his über-technician Dennis Sutton, I learned about *Agrobacterium tumefaciens* (wild-type strains and A66, a shooty mutant), and started to transform plant cells and culture the resulting crown gall cells. Referring to the race mentioned earlier, this was pre-disarmed *A. tumefaciens*, pre-gene gun, pre-GUS, pre-PCR, and so on, and there were limited standardized in vitro plant regeneration procedures developed for use. I'm glad we didn't know how bad we had it back then! I provided cell culture and plant transformation expertise that led to the first confirmed transformation of plant cells expressing a foreign eukaryotic gene (1983; phaseolin gene from bean introduced into sunflower; jokingly referred to as "sunbean"). Our company's next task was to develop transgenic plants containing, and hopefully expressing, phaseolin; the phaseolin gene contained its original developmentally regulated promoter for expression in seeds. Since, at the time, tobacco was much more regenerable than sunflower, I introduced the phaseolin gene into tobacco using A66 which could yield shoots, but never resulted in intact plantlets, so I taught myself how to graft tobacco shoots onto tomato rootstocks (worked better than

tobacco) to enable these shoots to grow and reproduce. Under Champa Sengupta-Gopalan's mentoring, I assisted with molecular analyses of transgenic tobacco and resultant seeds, and was proud to be a coauthor on the first paper that confirmed correct, developmental expression of a foreign eukaryotic gene in transgenic plants and Mendelian transmission to their progeny (1985).

These experiences confirmed my road would include graduate school with a focus on plant biotechnology. When Kemp left for New Mexico State University (NMSU) to start the Plant Genetic Engineering Lab, he asked me to be his graduate student, and since Sutton and Sengupta-Gopalan also accompanied him there, the move to NMSU felt right. Near the end of my dissertation research, I secured travel funds to attend the Horticulture Biotechnology Symposium held at UC-Davis in 1989, and presented a poster on my research. I also included a very small sign that stated "NR needs a job" in the bottom right corner. The head of the Department of Horticulture at Mississippi State University (MSU; shortly thereafter reorganized/renamed Department of Plant and Soil Sciences) attended that meeting and discussed a new tenure-track faculty position they would be advertising, and encouraged me to apply. I did, and have been at MSU for the past 25 years.

My research at MSU has primarily focused on plant tissue culture and genetic engineering, with a greater emphasis on developing and optimizing regeneration procedures for various plant species. Being mostly funded by the MS Agriculture and Forestry Experiment Station, the species I primarily worked on were those grown in the state (corn, cotton, kenaf, rice, soybean), and I received an internal science award for our publication on soybean regeneration. I enjoy conducting basic research but feel that it must also have an application in the near or distant future—which may result from my exposure to the biotechnology industry. Not forgetting about our students, some of whom will be our future scientists, I tried to contribute to their exposure to plant biotechnology through direct involvement in my lab and my courses: Plant Tissue Culture and Plant Genetic Engineering.

I diverged from my research path for 8 years while serving as Head of the Department of Biological Sciences at MSU. On the topic of educating future scientists, our department is responsible for educating 1100 undergraduate student majors each year, in addition to supporting courses for majors across campus. Critical to keeping the department vibrant to serve our students and science, I was able to hire a number of great faculty members, and it has been quite gratifying to observe and assist in their development into phenomenal scientists. Their contagious excitement and enthusiasm reminded me about myself before diverging into an administrative role—privileged to be able to conduct research that excited me and the freedom to explore, so I have since stepped back into the faculty. This has brought me full-circle and I plan to finish out my career in the lab, where I began. I am also leading a mentoring/retention program for female faculty members in the college, at the request of the dean, so I can continue to assist and contribute to faculty development.

I would also like to add that the scientific society I chose to get involved in, the Society for In Vitro Biology (SIVB, formerly called Tissue Culture Association), has also greatly influenced me; my involvement began at the recommendation of other important scientists in my life, Bob Lawrence (Agrigenetics Applied Genetics Lab in Boulder, Colorado) and Greg Phillips (NMSU). I have met a number of great scientists in this society, many who I am proud to call friends, and have had the great honor to serve in various leadership roles. In 2006, I received the society's Fellow Award. Going with the full-circle theme, I am also part of a once-yearly band of scientists (Hobbit Nirvana) that plays at our meeting's joint social, where I finally get to sing.

Robert Frost's poem, identified earlier, closes with the following (partial phrase provided): "Two roads diverged in a wood, and I—I took the one…" Divergences allow each of us to choose our road that may contain one primary divergence or multiple ones, with U-turns and circling back allowed.

REFERENCES

Harper BK, Mabon SA, Leffel SM, Halfhill MD, Richards HA, Moyer KA, Stewart CN Jr (1999): Green fluorescent protein in transgenic plants indicates the presence and expression of a second gene. *Nat Biotechnol* **17**:1125–1129.

Mentewab A, Stewart CN Jr (2005): Overexpression of an *Arabidopsis thaliana* ABC transporter confers kanamycin resistance to transgenic plants. *Nat Biotechnol* **23**:1177–1180.

Mullis KB (1990): Target amplification for DNA analysis by the polymerase chain reaction. *Ann Biol Clin* **48**:579–582.

Poovaiah CR, Mazarei M, Decker SR, Turner GB, Sykes RW, Davis MF, Stewart CN Jr (2015): Overexpression of switchgrass sucrose synthase (*PvSUS1*) increases plant biomass in transgenic switchgrass (*Panicum virgatum* L.). *Biotechnol J* **10**:552–563.

Southern EM (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**:503–517.

Yuan JS, Reed A, Chen F, Stewart CN Jr (2006): Statistical analysis of real-time PCR data. *BMC Bioinformatics* **7**:85.

Yuan JS, Burris J, Stewart NR, Mentewab A, Stewart CN Jr (2007): Statistical tools for transgene copy number estimation based on real-time PCR. *BMC Bioinformatics* **8**(Suppl 2):S6.

CHAPTER 13

Regulations and Biosafety

ALAN McHUGHEN

Department of Botany and Plant Sciences, University of California, Riverside, California

13.0. CHAPTER SUMMARY AND OBJECTIVES

13.0.1. Summary

Transgenic crops are the most regulated and tested plants ever produced, and much of the regulation is a response to concerns about biosafety issues. There are two areas of biosafety concerns: food safety and environmental safety, each with corresponding regulatory issues.

13.0.2. Discussion Questions

1. What are regulations supposed to achieve?
2. With GM crops spreading so quickly, how are we assured of their health and environmental safety?
3. How is agricultural genetic engineering (biotechnology) regulated?
4. How do the risks posed by products of biotechnology compare to those posed by conventional technologies?
5. How does biotechnology threaten biosafety?
6. How do different countries regulate products of biotechnology?
7. What are the salient differences between "process-based" and "product-based" safety regulations?

13.1. INTRODUCTION

This chapter explores how governments regulate food and agriculture emanating from one group of technologies, *genetic engineering* (also called *genetic modification, rDNA*, or simply "biotechnology"), and investigates the scientific validity of such regulations.

Our human ancestors began the serious art and science of agriculture about 10,000 years ago. In those days and until the near-present time, the major concern was simply getting enough food. Today's agriculture issues still include, for approximately 800 million people, getting enough to eat,

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

but also a range of other concerns, such as food safety and nutrition. In addition, other economic and political issues can occupy the minds of those who typically show few signs of hunger or malnutrition. The planet supports a burgeoning human population well over 7 billion; but without human intervention in genetics, nature can sustainably provide for only about 3–4 billion of us. For good or for bad, the success of humans at procreation now demands that we turn increasingly *against* nature in order to provide enough food to maintain the increasingly unnatural human population (McHughen 2014). As agriculture becomes increasingly technological, and less and less traditional, many people become increasingly vocal in expressing concern for safety in food production systems.

Our prior history shows little mass interest in the safety of food production, especially if there was sufficient safe food to go around. But societies have always suffered from local, regional, or widespread food famines and adulterations, and these scourges continue today. With public interest in food and agriculture increasing within affluent societies, newer technologies are coming under scrutiny as potentially hazardous.

The transition from traditional farming practices and food production systems to the application of modern technologies in all aspects of agriculture and food in the early twentieth century was accompanied by a mass exodus of farm folk to urban centers. In 1900, for example, agriculture occupied 41% of the US workforce; but by 2000, agriculture occupied less than 2% of US workers (USDA/ERS 2005). As a result, unlike a century ago, few urban people in affluent societies have a direct personal or family connection to farming and consequently have little comprehension of how food is produced. This unfortunate ignorance leads to gross misconceptions and a rather romantic aura of "traditional" farming. The anxiety fostered by beliefs that the agricultural technology is suspect also leads to demands that government assume a greater role in ensuring the safety and security of the food supply, even when there is little or no scientific justification (on the basis of actual harm) for doing so.

A large number of technologies—all of which pose some degree of risk to health or environment—have been introduced to farming and food production in the past 100 years. Many of these, such as mechanization, farm management (agronomy), and genetic modification through plant and animal breeding, have had a dramatic and positive impact on both the quantity and quality of food produced. In addition, technological advances and applications in food storage, processing, and transport allowed human society to eat, flourish, and expand well beyond natural limits to the sustainable population and allowed individuals to enjoy an expected average lifespan nearly double that of our grandparents.

Nevertheless, products of all technologies do carry risks, and in modern risk-averse society, those risks must be identified, assessed, and managed. Because of the long history of relatively "safe" introductions of technology to agriculture and food, most city dwellers paid little heed to risks associated with adoption of, for example, tractors on the farm, although many farmers (and family members) suffered death or dismemberment from mechanical accidents involving the powerful machines, and such accidents continue today. Through the twentieth century, governmental regulations evolved to ensure the safe application of almost all products and innovations in farming. However, in the 1970s and 1980s, many people began to question the safety of food production systems and the efficacy of regulations governing them. Spurring this anxiety, in the absence of any true problems with the food supply, was the increasing awareness and even fear of chemical fertilizers, pesticides, and the general feeling that farming was becoming "high tech," and not the way it was in the old days. One manifestation was a common wariness and subsequent demand to increase regulation on plant, animal, and microbial breeding, where genes were modified using recombinant DNA (rDNA) technologies, often called genetic engineering (GE) or genetic modification (GM) to produce genetically engineered/modified organisms (GEOs or GMOs). In response to this, governments around the world rushed to assure the public that "something was being done to protect the public and the environment from the hazards of genetic engineering" and establish regulatory mechanisms to oversee GE as applied to agriculture and food production.

13.2. HISTORY OF GENETIC ENGINEERING AND ITS REGULATION

Genetic engineering, recombinant DNA, is much older than most people realize. The first successful DNA "recombination" or human-mediated hybridization between two specific but diverse DNA strands was reported by Boyer and Cohen in 1973 (Cohen et al. 1973). At first, the scientific community itself recognized that the great power of the new technology also implied risk (Berg et al. 1974), and in 1975 a group of leading scientists convened at Asilomar, California, to discuss the issues. They called for a largely self-regulated set of guidelines to cautiously assess the risks with the emerging technologies. In the United States, the National Institutes of Health (NIH) in 1976 took the next step when it formalized and established strict rules to regulate rDNA research activities. Although the NIH guidelines applied only to federally funded rDNA research programs, many agencies (including the Environmental Protection Agency (EPA), Food and Drug Administration (FDA), and the US Department of Agriculture (USDA)) adopted the rules as sensible precautionary policy. The voluntary NIH guidelines thus became, in effect, mandatory for virtually all rDNA research conducted in the United States and internationally.

With the scientific community enthusiastic about the applications of rDNA and other forms of biotechnology, bureaucracies recognized the impending certitude that biotechnology would not remain an academic and laboratory novelty, and that manufacturers of products developed using the new technologies would eventually be seeking market and environmental release. Consequently, they began gearing up to deal with potential hazards. One of the first papers was from the Organization for Economic Cooperation and Development (OECD), which provided a standardized and workable definition of "biotechnology … the application of scientific and engineering principles to the processing of materials by biological agents to provide goods and services" (OECD 1982). Although the definition is unwieldy and captures virtually everything involving biological systems, including products of conventional breeding and food production systems, it remains widely used today and provides the basis for regulations in many countries. The OECD report also noted the necessity of regulating products of biotechnology, assuming that they, like everything else, were not inherently risk-free. By the mid-1980s, the living organisms generated as a result of rDNA research (also known as *transgenic organisms, genetically modified organisms* (GMOs), or *genetically engineered organisms* (GEOs)) were being generated and attracted attention because of their own potential for risk, particularly as potential threats to the environment and as food/feed safety hazards. In 1986, the US Office of Science and Technology Policy (OSTP) investigated the regulatory milieu and compiled a *Coordinated Framework for Regulation of Biotechnology*. This document coordinated the existing regulatory bureaucracy with relevant studies coming from the scientific community. They recommended adapting existing legislation and regulatory authority to encompass products of biotechnology, tapping existing regulatory expertise in relevant agencies, particularly the USDA, FDA, and EPA. Thus, GM plants would be regulated for food and feed safety concerns by regulators with appropriate expertise in FDA, those GM plants with pesticidal properties by EPA, and those with plant pest potential (environmental risks) by USDA. This coordinated effort and allocation of responsibility to different agencies continues today in the United States (Wozniak and McHughen 2012).

At about the same time, the OECD released a major study (based on its own recommendation in the earlier 1982 report) on biosafety related to biotechnology, often called simply the "Blue Book" (OECD 1986), which also remains widely quoted and cited today for its fundamental commonsense approach to risk assessment. It was the first scientific analysis to consider hazards that might be posed by transgenic organisms, and served as a standard from which many governments and regulatory agencies have based their procedures for assessing risks with products of biotechnology. It remains, even after 30 years, "fresh" in the sense that it was prescient, identifying legitimate risk concerns with rDNA technologies even before transgenic organisms were let loose on the environment.

In contrast, some other jurisdictions, notably those in the European Union (EU), believing rDNA to be so novel and potentially hazardous that existing legislation and regulatory expertise was not capable of handling it, created entirely new bureaucracies to regulate GMOs.

By the end of the 1980s, the US National Academy of Sciences issued a "white paper" declaring, among other things, that rDNA produced no new categories of risk, and that risk assessment should be based on the physical features of the product, not on the process by which it was developed (National Research Council (NRC) 1987). Subsequent studies from the National Academies of Science (via the NRC) on increasingly specific points dealing with risks posed by rDNA all came to the same general conclusion, that all methods of genetic manipulation can generate potentially hazardous products, that rDNA is not inherently hazardous, nor does rDNA categorically generate products posing greater risk than do other methods, and that risk assessment should focus on the features of the final product, regardless of the method of breeding (NRC 2000, 2002, 2004, 2010a).

Back at the lab, the techniques of gene splicing, as it has become known, have been applied to a wide range of products, including medical, industrial, and, yes, agriculture and food production. In the late 1970s, the early experimental successes saw genetically engineered microbes produce proteins from rDNA transferred genes, and the technical advances were quickly adapted to commercial applications, including generating human therapeutics. Human insulin produced by rDNA from the human gene transferred to bacteria was reported in 1978. This development led to the first approval for the first commercial application of rDNA technology, the diabetes drug insulin (trade name: Humulin™, from Genentech), in 1982. Many other pharmaceutical products developed using rDNA quickly followed.

Transgenic plants made their lab and greenhouse appearance in 1983, as three independent groups reported their developments at the Miami winter symposium, and other groups followed quickly.

In Belgium, Jeff Schell and Marc Van Montagu produced tobacco plants resistant to kanamycin and methotrexate (Herrera-Estrella et al. 1983; Schell et al. 1983). At Monsanto in St. Louis, USA, Robert Fraley, Stephen Rogers, and Robert Horsch generated transgenic petunia plants resistant to kanamycin (Fraley et al. 1983a, b). And in Wisconsin, John Kemp and Timothy Hall inserted a gene from beans into sunflower (Murai et al. 1983).

The first open-air field trials of transgenic plants were planted as early as 1985, but the numbers of trials, species, traits, and countries climbed dramatically in the late 1980s and early 1990s.

However, it took 10 years (to 1993) before the first whole plant was commercialized and grown unregulated in the field, a virus-resistant tobacco in China (Jia and Peng 2002; Macilwain 2003), followed by the first transgenic food crop, Flavr Savr tomato, in 1994. Neither GM product remains on the market today. The Flavr Savr failed because of inconsistent production capacity and delivery to market (Calgene, the company developing Flavr Savr, claims that they sold every tomato delivered to the stores, but that they were simply unable to keep up with demand); the Chinese tobacco was withdrawn because of pressure from smokers worldwide who feared that smoking GM tobacco (but not regular tobacco?) might pose a health risk.

The first GE food product, the milk coagulating agent chymosin, was developed in 1981 and, after various improvements, testing, and safety assessments, was approved and reached the market in 1988 (in the United Kingdom) and 1990 (in the United States). Most of the hard cheese now made uses this genetically engineered protein in place of rennet from calf stomach. Such cheeses are popular yet remain unlabeled, even in places where labeling based on the process of rDNA is mandated. Although only trace amounts of the enzyme remain in the final food product, it is disingenuous and misleading to consumers to claim that the cheese is "non-GMO," at least not without an explanation. More on this later.

The subsequent deployment and adoption of GE crop varieties has been impressive. According to ISAAA (James 2005), the one-billionth acre of commercial GE crop was grown in 2005, with the total acreage spread across 21 countries. In 2005 alone, according to James, GE crops covered 222 million acres (~90 million hectares (ha)) worldwide. This represents an impressive growth within an industry, namely agriculture, not known for quick adoption, particularly of controversial technologies. The major players remain fairly constant, with the United States, Argentina, Canada, and China leading the way, but also significant acreages in some smaller countries, including such

diverse lands as South Africa, Philippines, Iran, and Romania. Some 27 countries, including 19 developing countries, are cultivating GE crops (James 2013).

In the United States, the major GE crops include soybeans, corn, and cotton; biotech cultivars of these crops captured over 90% of their respective market acreages (USDA/ERS 2014). Minor commercialized GE crops include potato, tomato, and flax (all no longer grown; see Ryan and McHughen 2014), plus virus-resistant papaya and some squash (both remain in production). GE alfalfa, sugar beets, and plum pox-resistant plums have been approved, and GE crops currently under development for US farmers include disease-resistant citrus (to combat the devastating Huanglongbing disease, for which there is no resistance in the breeding germplasm (NRC 2010b), and a broad array of others. For a complete listing of US approved crops, see http://www.aphis.usda.gov/brs/not_reg.html.

Other crops in development but not (yet) commercialized include those more directly attractive to end-use consumers. In addition to the so-called Golden Rice and golden banana to treat vitamin A deficiency in poorer populations, we can look forward to non-browning low acrylamide potatoes, non-browning apples, high oleic soybean, as well as environmentally beneficial traits such as drought-tolerant corn and other crops.

Internationally, GE crops under development include improved versions of locally important crops, such as GE brinjal (aka eggplant or aubergine) in Bangladesh http://www.isaaa.org/kc/cropbiotechupdate/article/default.asp?ID=12550), and high-protein potato in India; corn in South Africa; broccoli; tomato; sweet potato; papaya; banana; winter melon; watermelon; rice; several tree events; and even transgenic animals (pigs) in Taiwan; rice, turfgrass, potato, and various local species of vegetables and produce in Korea; oil palm in Malaysia; and cassava in Kenya and other countries of east Africa. An exhaustive listing of GE species and traits in development around the world would be both extensive and quickly outdated. Those interested in the technical and regulatory progress of agricultural biotechnology in developing countries should consult www.isaaa.org frequently.

Most of the GE crops commercialized to date carry input traits such as disease or herbicide resistance, or pest control, but newer products are focused on output traits, such as enhanced nutritional profiles and removal of allergenic or other antinutritional proteins and substances. One reason why these "consumer-oriented" GE products are not available today is the long and expensive regulatory process. One point worth remembering is that all GE crop cultivars receive far more regulatory oversight and safety assessments than do similar crops with similar traits and therefore posing similar risks.

13.3. REGULATION OF GM PLANTS

Effective regulations protect the public and the environment from threats of harm. Also, because all regulatory bureaucracies have limited financial, human, and other resources, they must, in order to be effective, apply the regulatory maxim: the degree of regulatory scrutiny should be commensurate with the degree of risk posed.

Of course, regulatory bureaucracies, like all bureaucracies, do not always work as efficiently or effectively in practice as in theory. Political expediency often interferes with the strict adherence to the scientifically sound maxim. In addition, political expediency all too often trumps science. A fundamental and common example of such an error is the attempt to regulate a perceived hazardous *process*, when the scientific community emphatically states that *process* is immaterial, that hazards—when they exist—reside with *products*, regardless of the *process* used to make them (McHughen 2007).

A big problem in regulating based on a process instead of the features of a product is that scant bureaucratic resources (including expert personnel) are assigned on a fool's errand, looking for hazards that don't exist (experts around the world agree that rDNA presents no risks that are not already present in traditional agriculture). Meanwhile, with regulatory resources shunted away from

real threats—such as *Escherichia coli* or *Salmonella* in foods—leading to actual, predictable, and unnecessary harm.

Another problem with regulating processes instead of the potentially hazards products (based on their actual features) is that processes evolve, sometimes dramatically, such that the statutes or regulations designed to capture the products of earlier techniques no longer capture the potentially hazardous products from sufficiently evolved techniques. Countries that regulate agricultural biotechnology based on the process of rDNA, for example, are stymied when creative researchers modify the rDNA technique itself, circumventing the language in the rules and regulations. In this eventuality, policymakers, instead of recognizing the fundamental error and changing the rules to capture potentially hazardous products, instead hobble together amendments to extend the definition of rDNA to include the new techniques (e.g., see the definition of Genetic Engineering in Vermont's Act 120, which adds an entire category of non-GE processes to the definition of GE). Thus the policymakers and regulators are on a perpetual fool's mission to catch up to the technologies as they evolve beyond the definitions.

Several newer technologies—as well as slight twists to old technologies—test the rules and regulations governing agricultural biotechnology. Cisgenics (as opposed to transgenics) refers to the use of DNA sourced from within a species for transfer. For example, the rice bacterial blight resistance gene *Xa21* exists in the rice species, but not in cultivated rice varieties. The *Xa21* gene has been cloned and transferred into commercial cultivars using rDNA techniques (Wang et al. 1996). To some, the resulting genetically modified rice cultivar is not subject to GE regulation because it is not "transgenic" but rather "cisgenic." To others, the rice is indeed subject to regulation because it was developed using the process of rDNA. If the objective of regulation is to assure safety, then both are wrong; they should instead be asking the question of how the physical presence of the *Xa21* gene (or, more correctly, the resulting gene product and its metabolites because the gene itself, composed of DNA, is categorically regarded as "safe" along with all other DNA molecules) affects the safety of the new rice.

13.3.1. New Technologies

The major and fundamental problem with process triggered regulations is that the processes (techniques) keep changing. All major countries promulgated their regulations covering agricultural biotechnology in the 1980s to early 2000s. But genetic technologies keep changing, improving, adding new features, and discarding obsolete ones. In recent years, "traditional" rDNA (the process trigger for many regulations) is been rendered obsolete by advanced techniques such as RNA interference (RNAi), oligonucleotide-directed mutagenesis (ODM), zinc finger nuclease (ZFN) technology, clustered regularly interspaced short palindromic repeats (CRISPR), and RNA-dependent DNA methylation (RdDM) (for technical details on these, see Chapter 17). Other newer techniques, such as cisgenics, were developed after the regulations were promulgated and also raise uncertainty over regulatory coverage. In 2012, the European Commission sought scientific opinion from EFSA (European Food Safety Authority) on whether they needed to amend the GMO regulations to ensure capture of products of these new technologies (EFSA 2012a, b). Policymakers in other countries are scrambling to see if their own process—triggered rules and regulations, will capture these products. If they had only followed the scientific community's advice, and triggered regulatory scrutiny based on non-familiarity of actual features of a product, they wouldn't need to update their regulations so often. For example, the Plant with Novel Trait (PNT) system of Canada captures plants with unfamiliar features, regardless of the process used to acquire those features. With this product-based system, Canadian regulators are able to capture those things that actually carry higher risk, even if they're not genetically engineered. For example, great damage to ecosystems is caused by simple transport, introductions of new plants from other regions of the world, which then escape and become noxious weeds in the new territory (NRC 2002). Canada's PNT policy gives regulatory protection against this actual threat to the environment because the product features of the new plants can be assessed and controlled.

In addition to process-based statute or regulatory definitions failing to capture products the policymakers (likely) intended to capture, the same definitions do capture things the policymakers (likely) didn't mean to capture. For example, Article 2(2) of Directive 2001/18/EC defines a GMO as *an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination*. The Directive then goes on to list several exceptions (including wide cross-pollination and mutagenesis, illustrating the problem in having a process-based capture mechanism), but it fails to exempt things like ordinary (non-rDNA) tomatoes carrying *Mi* genes conferring nematode resistance. Nematode resistance in tomato cultivars traces back to Smith (1944), who applied embryo rescue *in vitro* to overcome the natural barriers to hybridization between two species (the *Mi* conferring nematode resistance gene does not exist in the natural *Solanum lycopersicon* germplasm). The *in vitro* embryo rescue technique clearly violates the European Directive, and so any tomatoes derived from Smith's work should be subject to EU regulation as a GMO. All tomatoes with nematode resistance derive from this embryo rescue hybridization conducted by Smith in 1944 (Ho et al. 1992). Tomato cultivars need not reveal whether or not they carry the trait, but approximately 20% of current tomato cultivars in the United States carry the nematode resistance genes according to a US tomato cultivar database (http:// cuke. hort.ncsu.edu/cucurbit/wehner/vegcult/tomatoai.html). Fortunately, US tomatoes are not subject to European regulatory oversight, unless they are exported into the EU.

Even those traditional breeding and agricultural practices like interspecific grafting, which humans have been doing for thousands of years, appear to be captured by European and other process-based regulations. Other non-rDNA techniques, including bridge crossing, translocation crossing, and somatic hybridization also appear to violate the European regulations (EFSA 2012a).

Other process-based definitions will restrict capture to the use of "*in vitro* rDNA" techniques, which captures most GE products, but is easily circumvented by a non-*in vitro* method to produce genetically engineered plants, such as the floral dip (e.g., Clough and Bent 1998) method. Obviously, potentially hazardous products could be generated from this (or other) methods, but process-based regulations will not catch them.

In the following text, we explore how some regulatory bureaucracies apply their allocated resources to agricultural biotechnology.

13.3.2. US Regulatory Agencies and Regulations

Discussion of regulatory policy for products of biotechnology in the United States started relatively early. As mentioned earlier, the Office of Science and Technology Policy (OSTP), recognizing that potential risks and regulatory expertise were distributed across several bureaucracies, developed a coordinated framework to assign responsibilities to those relevant agencies (OSTP 1986). Within this, *regulated articles* (as they are called) were assigned to the different agencies according to their intended use, but also recognizing that some articles—and, in practice, most—were captured for regulation by more than one agency. As a result, FDA was given primary responsibility for regulating risks to food and feed, EPA to regulating products with pesticidal properties, and USDA to biotechnologically derived plants with potential to become agricultural pests. In many cases, all three agencies evaluate a product; for example, a food crop with rDNA-mediated novel herbicide resistance would trigger review by USDA for plant pest potential, EPA for the new herbicide aspects, and FDA for any changes to the quality of the derived food and feed.

Other products, for example, an ornamental (nonfood/feed) plant with an altered flower color, might avoid regulatory review by EPA and FDA, but still be captured by USDA. In fact, until recently all rDNA plants seeking deregulation were captured and regulated by USDA. All commercialized rDNA-derived food crops were reviewed by FDA, even though the food itself was unchanged and thus the FDA assessment was considered "voluntary" (much to the dismay of some, who believe that FDA should regulate all biotech products as a mandatory exercise).

The United States conducts regular evaluations of its own regulatory procedures, to ensure that the regulators remain aware of the most recent developments in the technology, and may adapt regulatory procedures to account for those developments. The scientific foundations are often reviewed by committees ("panels") of the National Research Council under the administration of the National Academies of Science. Administrative procedures are also frequently reviewed, usually involving solicitation of public input and suggestions for improvement. In addition, public input is sought at several stages of the regulatory review, usually after an announcement in the *Federal Register* detailing a particular product under review. For more detail on how the US regulated products of agricultural biotechnology, see Wozniak and McHughen (2012).

13.3.2.1. United States Department of Agriculture. The office within USDA responsible for regulatory oversight of agricultural products of rDNA is the Animal and Plant Health Inspection Service (APHIS), office of Biotechnology Regulatory Service (BRS). Regulators in BRS claim legislative authority to capture and regulate rDNA-derived plants under the Plant Protection Act of 2000. The main concern in BRS is that the "regulated article" (i.e., rDNA-derived plant) might become a "plant pest" (defined broadly) and negatively impact the environment, so they focus their assessments on pest characteristics. USDA assesses whether the regulated article (product of rDNA breeding) might directly or indirectly cause disease or other damage to a plant. The primary regulated articles to date are herbicide-tolerant crops, insect-protected crops, and a handful of other transgenic plants (virus-resistant squash, herbicide-tolerant sugar beet and alfalfa, late-ripening tomatoes, virus-resistant papaya, potatoes, etc.), as well as some transgenic microbes. BRS controls not only prospective releases to the open environment but also the international importation and interstate transport of transgenic organisms.

BRS allows environmental releases of transgenic plants through two routes: notification and permit. Notifications are used for specified low-risk crops and traits, while a permit is required for those transgenic organisms posing greater apparent risk, such as species less familiar to BRS or those producing pharmaceutical compounds. Eventually, after the evaluations are complete; and if the data support it, the developer may petition for "unregulated" status. BRS conducts an environmental assessment to ensure that the product is indeed not a potential plant pest, and also seeks public comment before issuing the decision. Once a "regulated article" acquires "nonregulated" status, it can be grown, sold, and distributed much like any other nontransgenic variety.

13.3.2.2. Food and Drug Administration. FDA, which operates within the Department of Health and Human Services (HHS), concerns itself with the safety of foods and feeds. Interestingly, unlike the case in USDA, in FDA the trigger for mandatory capture for regulatory assessment is not the process of rDNA, but the physical composition of the food or feed in question. This is the basis of considerable debate, as some people demand that FDA conduct safety assessments of all foods derived from biotech plants, animals, and microbes, even those with chemical compositions identical to those of current foods of the same type.

The FDA review focuses on three questions:

1. Does the novel food or feed contain any new allergens?
2. Does the novel food or feed contain any new toxic substances?
3. Has the novel food or feed changed the nutritional composition in any way, either increasing or decreasing nutrients, antinutritional substances, or other components?

Problems from ingesting food result from the presence of damaging substances such as allergens in sensitive people, or toxicants. In the long term, problems can also arise from the absence or diminution of nutrients ordinarily present in a given food. For example, many people enjoy orange juice and benefit from the rich source of vitamin C. If for some transgenic reason oranges ceased to produce ascorbic acid (vitamin C), some consumers might inadvertently develop symptoms of

vitamin C deficiency, namely, scurvy. As this is an undesirable effect, FDA would check a new orange for ascorbic acid content, just to ensure that it was still present in appropriate concentrations. To date, biotech-derived (biotechnologically derived) foods have not been found to unexpectedly lose normal nutrients, and all commercialized biotech-derived foods have the same nutritional content as do similar conventional foods. Newer transgenic foods might be specifically modified to enhance nutritional composition. In those cases, the FDA review becomes mandatory, and the new food will have to be labeled as such, because it would no longer fit the definition of the traditional, unmodified food.

A bigger concern is the possibility that the novel food carries an unexpected allergen. Such an event has occurred, although the product was never commercialized and no one was harmed. In this situation, a gene to enhance the nutritional status of soybean (which is naturally deficient on the amino acids methionine and cysteine) was cloned from Brazil nut and transferred to the legume. Tests showed that the transgenic soybean did indeed express the Brazil nut gene and generate the expected protein rich in these amino acids, thus successfully increasing the nutritional balance of the bean. Subsequent premarket tests showed that the new soybean was, unfortunately, also allergenic to consumers allergic to Brazil nut, indicating that the storage protein in Brazil nut responsible for the good desired amino acids was also a major allergen, even when expressed in soybean (Nordlee et al. 1996). Since the Brazil nut transgenic soybean was found to be a likely source of allergens during the course of evaluation, it is heralded as a case showing that regulations are effective.

Even without a mandatory premarket food safety assessment, every commercialized rDNA crop was reviewed by FDA regulators under a voluntary consultation. In other words, the developers of the new crops and foods wanted the FDA to review the safety even though it was not legally required. The reasons are clear enough; developers want help from FDA to ensure that their new products are safe before putting them on the market. Without that safety check, a new food released onto the market and later found to have, for example, new toxic substances would face (1) regulatory action from FDA for releasing an adulterated food and (2) litigation from unsuspecting consumers harmed from ingesting the adulterated food. With the dire consequences, especially of the latter, and with the simple and sensible procedures in the "voluntary" FDA consultation, any biotech food developer who bypassed the FDA review would be nothing short of foolhardy.

13.3.2.3. Environmental Protection Agency. The US EPA is concerned with risks posed by pesticides (including herbicides). According to EPA, a pesticide can be any substance or combination of substances intended to prevent damage by any pest, or intended for use as a plant growth regulator. For transgenic plants, this usually means herbicide-tolerant or insect-protected cultivars, but can include others also, especially those the agency calls "PiPs," for "*Plant incorporated Protectants*." Importantly, EPA claims that it does not regulate the transgenic plant *per se*, but rather they regulate any pesticidal properties associated with the transgenic plant. Because of this pesticidal properties trigger, not all transgenic plants require EPA regulatory approval. For those transgenic plants with pesticidal properties, EPA issues permit for large-scale (>10-acre) field trials and seed increase plots, and also regulate commercial registration for any such plant varieties sold with pesticidal claims, such as *Bt* corn or herbicide-tolerant soybeans.

13.3.3. European Union

The EU seems most confused on the issue of biotechnology. Many leading scientific technical developments in biotechnology have occurred within the borders of EU nations, but the application and deployment of the technologies is chaotically skewed, with seemingly rapid commercialization of medical, food, and industrial biotech applications, while lagging in GM crop approvals and releases. It seems contradictory that hard cheese in the EU emanating from a GMO, albeit lacking in detectable GE DNA or protein, is exempt from regulations or special labeling. But corn, canola, or soybean oil, similarly lacking detectable GE DNA or protein, is so captured for both extensive

regulatory oversight and product labeling. Internally, the EU is not so united with regard to their views of regulation. Several member states appear at least somewhat supportive of agricultural applications of biotechnology, others are more hesitant, and several remain rigidly hostile. In 2014, Europe moved to give member states more independent authority to restrict or even ban the cultivation of GMOs, even those fully approved throughout the EU under Directive 2001/18 (http://eur-lex.europa.eu/resource.html?uri=cellar:303dd4fa-07a8-4d20-86a8-0baaf0518d22.0004.02/DOC_1&format=PDF). One intent is to give more open-minded countries (such as the United Kingdom) the opportunity to grow EU-approved GM crops. But, as that intent is not written into the rule, whether or not it occurs remains to be seen.

Regulations are split among several pieces of legislation. In the early days of agricultural biotechnology, the EU split their regulations between two regulatory Directives: 90/119/EC covered contained use of genetically modified microorganisms, and 90/220/EC, which covered deliberate environmental release of GMOs. Both of these were later substantially amended; 90/119/EC was superseded in 1998 by 98/81/EC, and 90/220/EC was superseded by 2001/18/EC in 2001. In addition, Regulation (EC) No. 1830/2003 amended Directive 2001/18/EC, outlining traceability and labeling provisions for GMOs and their derived foodstuffs. Regulation EC 1829/2003 provides specific details for labeling requirements.

A listing of the GMOs that are authorized in the EU is available at http://ec.europa.eu/food/dyna/gm_register/index_en.cfm.

In addition to these primary regulatory documents, Regulation EC 258/97, superseded by EC 1829/2003, covers approvals for "food and feed consisting of, containing or produced from genetically modified organisms," and Regulation EC 1946/2003 http://ec.europa.eu/food/plant/gmo/trans-boundary_movement/index_en.htm) provides the EU procedures governing the transboundary movements (i.e., international trade) of GMOs, effectively implementing the Cartagena Protocol, as well as unintentional transboundary movements.

Complicating this already complicated bureaucracy is the "safeguard clause," which allows member states to essentially opt out of accepting GMOs deemed safe under the various regulatory directives. This escape clause has been used liberally by member states hostile to GMOs. Member states invoking the safeguard clause are required to submit scientifically sound justification for rejecting the determination of safety; but in every case, the scientific committee—and sometimes the courts—failed to find scientifically valid justification. Despite this, conflict within the vast European bureaucracy, GMOs remain relatively scarce in the farmers' fields (to date, although a few member states—Czech Republic, Slovakia, Romania, Portugal, and Spain—are cultivating a single GMO—a Bt Maize—and those are on tightly limited acreages). Foods derived from GMOs are even rarer, except for such examples as the hard cheeses produced with enzymes from GMOs, which escape regulatory scrutiny and labeling due to a convenient semantic distinction between foods produced *from* GMOs (which are captured for regulatory scrutiny) and those foods produced *with* GMOs (which, like the cheeses, are curiously exempt). Considering the public anxiety in Europe surrounding GMOs, it seems odd that the general public would appreciate the distinction between *from* and *with* to the extent that a food made *from* GMOs faces a heavy regulatory burden while a similar one, posing similar (insignificant) risk, made *with* GMOs gets a free pass without so much as a label. Paradoxically, EU consumers consume a broad range of imported GMO food ingredients, including vegetable oils from GM maize, canola, soybeans, and cottonseed, along with sugar from GM sugar beets and meat from European animals fed on imported GM feeds. The net result here is that the EU farmers are denied access to GM crops used by their competitors overseas, putting them at disadvantage with their own local consumers in the marketplace.

Unfortunately, although the EU policies all claim to stem for a concern to protect health and the environment from risks associated with GMOs, nowhere are such risks documented and ascribed to GMOs specifically; and if there really are actual risks, the EU consumer is exposed to them via the large amount of imported GM foods and feeds. Nevertheless, the EU strictly regulates almost all

aspects of agricultural biotechnology and resulting products, making EU-approved products the most scrutinized products ever to reach the commercial marketplace.

Interestingly, the United States, Canada, and Argentina brought suit against the EU in the World Trade Organization (WTO), claiming that these regulatory measures were illegal because they appeared to discriminate against "foreign" products of biotechnology and served as an illegal trade barrier. The WTO agreed; but the final resolution, if there is one, will probably take several more years. A major issue is the focus on the assumed risks posed by biotechnology and its products. With scientific studies worldwide unable to document any health or environmental risks unique to GMOs, the EU was hard-pressed to justify their position in establishing regulations to protect against health and environment against "the risks inherent in GMOs." Indeed, European scientists have been actively busy searching for such risks for several years. According to Kessler and Economidis (2001), the European Commission itself spent 70 million Euros to fund 81 research projects employing 400 teams of scientists between 1984 and 2000 to characterize risks associated with GMOs. A second document was issued in 2010, covering the EU sponsored research on GMO safety for the previous decade, this one consisting of 50 new projects conducted by 400 public research teams spending 200 million Euros (http://ec.europa.eu/research/biosociety/pdf/a_decade_of_eu-funded_gmo_research.pdf). In total, the European Union spent €270,000,000 over 25 years seeking a scientific rationale for justifying their discrimination against GM breeding methods. But not one risk unique to GMOs was found.

13.3.4. Canada

Canada remains unique worldwide for recognizing that risk is posed by potentially hazardous products, not by the process by which the products are made, and captures for regulatory oversight "novel" products, even some not developed using rDNA or other forms of biotechnology. Currently, all other jurisdictions use a process-based trigger for regulatory capture, and that process is rDNA (although the legal definitions of "rDNA" and "biotechnology" do vary considerably). To date, Canada remains the only country where the conclusions of the scientific community (viz., that breeding process is unrelated to risk) have been adopted into the regulatory practice. Once regulatory action is triggered, however, differing jurisdictions are remarkably consistent in their scientific risk assessments.

Canada assigns regulatory responsibility to three main federal agencies: the Canadian Food Inspection Agency (CFIA), Health Canada, and the Canadian Environmental Protection Agency (CEPA). Health Canada is responsible for food safety exclusively, while offices within CFIA are concerned with environmental issues and threats to animal feed. CEPA provides an insurance "catchall," capturing anything that appears to "fall through the cracks" or find "loopholes" to ensure nothing avoids regulatory scrutiny altogether. The relevant Health Canada website for "novel foods" is here: http://www.hc-sc.gc.ca/sr-sr/biotech/food-aliment/index-eng.php#app. The "novel plant" approvals by CFIA in Canada can be found at http://www.inspection.gc.ca/active/eng/plaveg/bio/pntvcne.asp, and the explanation for the approval process can be found at http://www.inspection.gc.ca/plants/plants-with-novel-traits/general-public/assessment-process/eng/1338189630096/1338189929476.

13.3.5. International Perspectives

Regulatory agencies worldwide recognize that products of biotechnology can pose risks, the same as can similar products from other means of genetic manipulation, including traditional breeding. Simply because they are generated using rDNA does not make them benign; they may have food or feed safety issues, and they may have features enabling them to threaten ecosystems.

Food safety is a common fear, and food safety regulatory agencies worldwide consider the possibility that the regular food has intentionally or unintentionally become adulterated, toxic,

or allergenic during the breeding process, or has significantly reduced (or enhanced) nutrients. All such agencies question the source of an introduced gene, to determine whether, for example, that source is allergenic. Because of the earlier work showing the allergenic Brazil nut storage protein to be allergenic even after gene transfer and expression in soybean (Nordlee et al. 1996), we know that allergenic proteins do not need their "home" genetic or physiological background to elicit an allergenic response. Similarly, food safety agencies are concerned with the potential for the transfer of potential toxic and other antinutritional substances from donor species, and the possibility that the transfer of even benign genes and proteins might exacerbate production of endogenous toxins, allergens, and antinutritional substances in the recipient species and foodstuffs.

Fortunately, both traditional and biotech crop and food developers also appreciate these real risks and conduct premarket testing to convince themselves (if not everyone else) that their new variety carries no additional toxic, allergenic, or antinutritional substances. Any breeding lines exhibiting such problematic substances are eliminated from consideration for commercialization long before any regulatory agency sees them. No company wishes to face the liability of releasing a true threat to health.

Also, fortunately, human physiology being what it is, a toxin, allergen, or other antinutritional substance will pose the same risk to virtually all populations worldwide. While there may be some differences in exposure, due to cultural or cuisine preferences or preparation methods, a toxin to western Europeans will also be toxic to Indians. Potential allergens might only elicit a response in a fraction of the population, and therefore any protein that contains a stretch of amino acids that cause allergens in anyone should be avoided. Nonetheless, all this means that the basic safety testing will be common to all, so the questions asked and answers demanded by the US FDA or EFSA in Europe will be of interest to consumers worldwide, and food safety regulators need not duplicate the entire (and expensive) food safety bureaucracy, but may instead concentrate on local variations in cuisine, including consideration for method of preparation (e.g., cooked vs. raw) or overall exposure (e.g., a food may be a major dietary component in one culture and minor elsewhere).

The other main scientific concern for regulatory action, environmental risks, is more variable. Consensus in the scientific (if not always in the regulatory or political) community recognizes that the factors in environmental risk are not the method of breeding but the species in question, the trait, and the region of release. While human physiology is much the same worldwide, ecosystems vary widely, such that a plant-deemed benign by USDA APHIS BRS for release in the United States might be wreak ecological havoc when released in the Amazon basin. Because of the environmental variation, regulatory agencies worldwide concerned with ecological effects cannot rely entirely on determinations made by regulators in a different environmental region. As not all countries or regions enjoy the regulatory resources of the United States, Canada, or the European Union, international efforts and regional coalitions attempt to economize biosafety review of potential threats to the environment.

One predominantly scientific society devoted to assessing environmental risk from products of biotechnology is the International Society for Biosafety Research (ISBR), which holds biennial symposia to discuss various scientific and regulatory developments concerning biosafety and how biotechnology may affect the biosphere, and is particularly concerned with the issues as they relate to developing countries. The proceedings of the last several years of these symposia are available online at http://www.ISBR.info. The ISBR site also includes links to several GMO approval databases.

Another attempt to consolidate information on the risks of products of biotechnology and their potential effect on biodiversity, particularly in poorer countries, is the Cartagena Protocol, which emanates from the Convention on Biological Diversity (CBD); see http://bch.cbd.int/protocol/. The objective is "to protect biological diversity from the potential risks posed by living modified organisms resulting from modern biotechnology" (http://bch.cbd.int/protocol/background/). Over 165 countries have signed the protocol, which obligates signatories to establish bureaucracies to identify,

monitor, document, and track living modified organisms (LMOs). The agreement covers international trade of the designated LMOs, which means viable, nonprocessed products of biotechnology. Essentially, this means grains and oilseeds such as soybeans, maize, canola, and cottonseeds, but not vegetable oils or food products derived from the commodities. Important to note, however, is that major international grain exporters including the United States, Canada, Argentina, and Australia are not members of the Protocol and are not bound by its provisions.

One useful provision of the Protocol is the Biotechnology Clearing House, a repository of information on living modified organisms (which, unfortunately, is defined by the process of biotechnology, not to actual threats to the environment). The Protocol, now ratified by 134 countries (although, to date, no major agricultural exporters, including the United States), and the clearinghouse database allow countries access to information on particular GM crops and assist in making regulatory decisions on the degree of risk to local ecosystems. The portal to the clearinghouse is available online at http://bch.cbd.int/database/organisms/ (simple, free signup is required to gain access). Unfortunately, the Cartagena Protocol is founded on the assumption that products of biotechnology present a threat to biodiversity (see previous paragraph), but no evidence to support this assumption exists. Particularly unfortunate is the corollary assumption that biodiversity is threatened *only* by biotechnology, as all non-LMO grains, oilseeds, and other viable commodities in international trade are exempt. The many scientific studies of environmental risks posed by products of biotechnology invariably conclude that products of biotechnology do not pose any greater threat to environment than do conventional products, thus invalidating the underlying assumption of Cartagena Protocol. There remains not a single documented case where a GMO (or LMO) has caused harm to biodiversity (McHughen 2006). This means that the true threats to biodiversity, the things that have wreaked havoc in our planetary ecosystems over the years, from introductions of invasive species to monocultures of popular (non-GMO) crop genotypes, will continue unabated, because Cartagena directs all regulatory resources to protecting against hypothetical risks (in LMOs) and no effort to stop the things that actually cause real harm.

13.4. REGULATORY FLAWS AND INVALID ASSUMPTIONS

Most current regulatory systems are scientifically flawed, despite assertions from politicians and regulators claiming that their system is indeed "scientifically sound." There are several scientific flaws, and any one of them invalidates the entire regulatory structure. Most importantly, if the objective is to assure food and environmental safety, regulations will have to change from a "process trigger" to a product trigger. That is, in most jurisdictions, safety regulations are applied only to those new plants developed using a process of rDNA, while potentially hazardous new plants developed using "traditional" breeding processes—or even introductions from elsewhere—are explicitly or implicitly exempt. This state of regulatory trigger is scientifically unjustified, practically inadequate (due to the triggering processes constantly evolving) and potentially dangerous for at least two reasons:

1. True hazards are present in products, regardless of the process used to create them. By focusing entirely on a breeding process (i.e., the use of rDNA), regulators are assigning regulatory resources to regulate a process the world scientific community has declared non-hazardous since the mid-1980s; and in so doing, they overlook the actual threats from products with potentially hazardous features, regardless of the process used in their development.
2. The cost of regulatory compliance is so great (estimated in the multiple millions of dollars, with no concomitant increase in safety assurance) that it incentivizes regulatory circumvention by developers who seek processes that are not captured by the legal language in the regulations.

For example, strict USDA language on "plant pest" led to development of methods and products to skirt regulatory review; for example, Scott's developed a GE Kentucky Bluegrass variety that was declared exempt from USDA regulations because it carries no plant pest components (https://www.aphis.usda.gov/biotechnology/downloads/reg_loi/scotts_0512_kbg.pdf). (Note, this does not mean the GE turfgrass is completely unregulated, as EPA approval is still likely required because of the pesticidal properties.)

In addition to being potentially dangerous, the process-based regulatory trigger is counterproductive. Not only does it stifle innovation (many public benefit GMOs are stuck in university storehouses because universities and other public institutions cannot afford the high price of regulatory compliance, and may public sector scientists have ceased developing such GM products (Miller and Bradford 2010), it also denies or delays the benefits of such products to consumers and the environment. Society might accept these foregone benefits if there were a rational explanation, such as a documented and unmanageable hazard. But there is none.

Six common invalid assumptions in regulatory policies are listed and discussed in the following text.

13.4.1. Conventional Plant Breeding has Higher Safety than Biotechnology-Derived GM

Most regulatory bureaucracies assume that "traditional" means of genetic modification are risk-free, while the processes of biotechnology inherently pose risk (Fig. 13.1). This assumption is rarely challenged, despite scientific studies over several years and from many countries establishing that the processes of biotechnology are not inherently more hazardous than other breeding methods (see, e.g., OECD (1986), NRC (1987, 2004), and Kessler and Economidis (2001) and http://ec.europa.eu/research/biosociety/pdf/a_decade_of_eu-funded_gmo_research.pdf, 2010). Indeed, in over 25 years of active research and experience there is still not a single documented case of harm from the consumption or cultivation of GMOs not also seem with products of traditional breeding. Obviously, accepting that rDNA poses the same risks as traditional breeding opens the door to potential regulation for *all* products of plant breeding, not just those derived from rDNA. And since there is little or no public demand to launch risk assessments for conventional agriculture, the only scientifically valid position is to relax the strict regulation of at least some benign GMOs to the level of that imposed on conventional agricultural products of similar risk. In many parts of the world, relaxing regulatory oversight of GMOs is politically unpalatable, even if scientifically justified.

13.4.2. GMOs Should Be Regulated Because They're GMOs and Un-natural

A major motivation in some jurisdictions to regulate GMOs exclusively is the assumption that transferring genes across the species barrier is unnatural and potentially hazardous. However, the concept of a rigid species barrier is itself inherently flawed, as there are countless examples, both in nature and under human manipulation, of moving genes from one species to another without added hazard. *Agrobacterium tumefaciens* is a natural genetic engineer, moving short pieces of prokaryotic bacterial DNA into eukaryotic nuclei and having the transferred genes integrate into the host genome. Many popular bread wheat (*Triticum*) cultivars carry fragments of rye (*Secale*) chromosomes.

13.4.3. Even though Product Risk is Important, It is Reasonable that Process (GMO) Should Trigger Regulation

No jurisdiction has sufficient resources to "test everything for everything," so a sensible system to prioritize regulatory resources evolved the maxim that products posing the greatest risk should face the greatest regulatory scrutiny. But this sensible approach has been abandoned in the case of

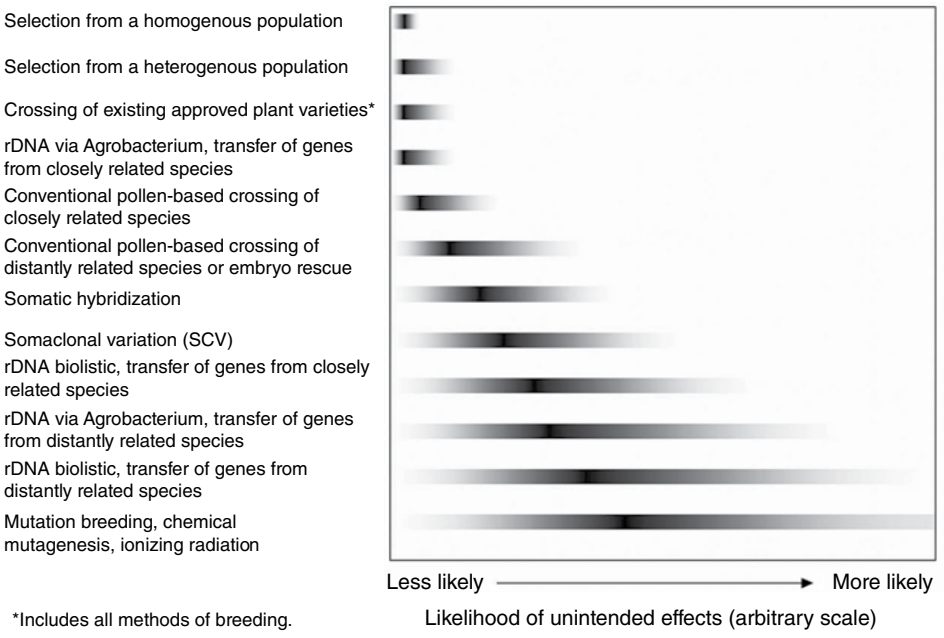


Figure 13.1. Relative likelihood of unintended genetic effects associated with various methods of plant genetic modification. The gray tails indicate the conclusions about the relative degree of the range of potential unintended changes; the dark bars indicate the relative degree of genetic disruption for each method. It is unlikely that all methods of genetic engineering, genetic modification, or conventional breeding will have equal probability of resulting in unintended changes. Therefore, it is the final product of a given modification, rather than the modification method or process, that is more likely to result in an unintended adverse effect. For example, of the methods shown, a selection from a homogenous population is least likely to express unintended effects, and the range of those that do appear is quite limited. In contrast, induced mutagenesis is the most genetically disruptive and, consequently, most likely to display unintended effects from the widest potential range of phenotypic effects. (*Source*: Reproduced with permission from National Research Council (2004).)

biotechnology. Consider two canola cultivars, one made resistant to acetolactate synthase (ALS) inhibitor herbicides using rDNA, and the other with identical herbicide resistance, except that it was developed using induced mutagenesis. The two similar cultivars pose similar risks, yet the biotech cultivar faces far greater regulatory scrutiny. Similar cultivars should face similar degrees of regulatory scrutiny, because they pose the same risks.

13.4.4. Since GM Technology is New, It Might Be Hazardous and Should Be Regulated

There is an unsubstantiated assumption that the risks posed by biotechnology are unique and should be evaluated as absolutes. However, risk is relative or comparative. Instead of asking "What are the risks associated with this GM crop cultivar?" a scientifically valid question is "What are the risks associated with this GM crop cultivar relative to the risks associated with the conventional cultivar that it will displace?" By focusing exclusively on the "new" thing and ignoring the status quo or current counterpart cultivar, any identified risk with the GM cultivar (and everything poses some degree of risk) can be and has been used as an argument to justify

banning the GMO, even though a proper, relative risk assessment might show it to be substantially superior to the riskier but currently grown cultivar.

13.4.5. If We Have a Valid Scientific Test, Then It Should Be Used in Regulations

The assertions that regulations are scientifically sound are invariably buttressed by scientific documentation showing the technical validity of the various assays, tests, measurements, and other criteria required by the risk assessors. But this line of argument merely supports the technical, not the overall, scientific validity. Technical skill in conducting a technically sound assay is not sufficient to satisfy scientific validity; it is necessary in addition to scientifically justify the rationale for conducting the test in the first place. For example, testing an extracted purified protein from a GM plant for possible allergenicity might seem a prudent regulatory requirement. But conducting the allergenicity trials is not scientifically valid, even if the trials themselves are conducted in a technically sound manner, unless there is a hypothesis or evidence suggesting that the protein may actually be allergenic. If the genes were cloned from a known allergenic source, or if the protein shared amino acid sequence homology with a known allergen, then yes, the technical allergenicity assays might be scientifically valid and prudently required. But to demand and conduct such trials merely to show the public that scientific tests for potential allergenicity were being conducted, or to exercise control over the developer, fails to increase real confidence in the safety of the product and jeopardizes public trust when the test was later found to be unnecessary, done only to appease public concerns.

13.4.6. Better Safe than Sorry: Overregulation is Better than Underregulation

Most risk assessments of GM plants are overly onerous and unnecessary in terms of informing risk management policies. Once sufficient data are collected to reach a determination of relative safety (or otherwise) of the GM cultivar, the law of diminishing returns kicks in; the cost of input resources escalate dramatically, but the additional data gleaned from the expenditure are usually superfluous and unconstructive. These additional data requirements undermine public confidence without adding any compensating features or increased assurance of product safety. The curious public wonders why so many additional and apparently unnecessary tests are demanded, and speculates that perhaps this product really is more hazardous than the developer and regulators are letting on. So the barrage of demanded tests and assays, instead of increasing public confidence, has the opposite effect—the public becomes even more suspicious and distrustful of both the product and the regulatory system.

The lesson is simple; to increase public trust as well as to increase confidence in product safety, before imposing and requiring any test or assay, the regulator should be able to answer "How will the information from this test/assay help inform or increase confidence in the safety (or otherwise) of this product?" If the answer is simply "more of the same," the regulatory demand is not scientifically valid and disrespects the public right to effective regulation.

Health is sometimes threatened by food-borne hazards (BSE, dioxins, *Salmonella*, and foot-and mouth disease outbreaks in EU; diarrhea-causing strains of *E. coli* in organic produce in the United States; etc.). Ecosystems and biodiversity have clearly suffered from human agricultural activity and breeding, such as the introductions of invasive species, particularly in Australia and North America. To date, there are no verified cases of damage to human or animal health, or to the environment, from GMOs (NRC 2004; AAAS 2013). All recorded harms come from non-GMOs. Yet almost all regulations capture for scrutiny only those products resulting from the process of biotechnology, and explicitly exempt non-GMOs, the sources of all known damage. The disconnect between the regulatory practice and the scientific recommendations ensures continued threats and damage to health and environment, and will do so until regulations capture and scrutinize those products posing the greatest risk, regardless of breeding method.

13.5. CONCLUSION

And so we come full circle. At the start of the chapter, we noted that regulations are supposed to protect us and our environment from various threats. We know the most effective way to assure safety is to identify the risks associated with a given product and assess those risks by comparing the features with the closest familiar object. Then, applying the regulatory maxim of the degree of scrutiny is commensurate to the degree of risk, we prioritize resources such that higher risk things face greater safety scrutiny, while lower risk products face lesser scrutiny. All the while, we learn from past experience and adjust the degree of scrutiny based on increasing familiarity of the category of product. Most importantly, regulators must focus on things, not on the process by which the things were made.

The main problem with the process trigger for regulatory scrutiny—as most jurisdictions do for biotechnology—is that we misappropriate resources, thus allowing the things that actually cause harm to be under regulated, or even escape scrutiny altogether.

It also means we spend too much time reviewing regulations as processes evolve. For example, with regulation of GE crops and foods, the EU established a process trigger to capture and regulate the safety of GE based on a definition of rDNA and associated processes in 1990. In the intervening years, the technology has evolved to add new techniques and modify the old, such that the regulatory capture of newer technologies—and even some older ones (grafting, embryo rescue), is uncertain. Indeed, the European Commission has had to commission a series of scientific studies in seeking guidance on whether products from uncertain processes (techniques) would be captured (EFSA 2012a, b).

In focusing in the process as a trigger for safety regulation, the whole exercise has gone from safety to semantics, with scarce regulatory safety resources being squandered in attempting to keep up with evolving processes, instead of being used to assure actual safety of all unfamiliar products (as is done in a product-based regulatory approach focused on familiarity, or novelty).

The product-based familiarity approach assesses new products and the differences between what we already have comfort with, so that a safety assessment of a new type of crop cultivar, a Bt corn, for example, would focus on the features of the corn (compared to a familiar, preferably isogenic corn line) combined with the Bt (e.g., the Bt we've used to control insects since the 1940s) and any novel features arising from the combination of the two. In this manner, we direct resources specifically at the potential threat. This science-based regulatory approach also precludes endless revisions to the regulatory language in an attempt to keep up with new processes, as the process by which a new or unfamiliar threat arises is immaterial.

The United States assessed and approved the first Bt corn after a massive regulatory assessment in 1995–1996. Using a process-based regulatory trigger, every Bt corn subsequently had to undergo the same risk assessment procedure. To date, regulators have assessed over 40 Bt corn events as if each were the first one submitted. But none of these showed any indication of being any more of a threat than the first one, so regulatory time and resources were effectively wasted in reviewing products where the risks were already known.

If the regulators had been operating on a scientifically sound product approach, the first Bt corn will have provided everything needed to know about the safety of corn genetically combined with Bt; the regulatory scrutiny on the second and subsequent Bt corn varieties would have been based on the differences between the first Bt corn and the second (or subsequent), thus making greater use of regulatory resources while still providing reasonable assurance of product safety.

As Dominic Grieve, the UK MP and former Attorney General said recently, "…one of the principal responsibilities of the government is to safeguard its citizens" (http://www.theguardian.com/commentisfree/2014/sep/03/terrorism-passports-rule-of-law?CMP=fb_gu). By adopting process-triggered safety regulations (whether concerning GMOs or other), governments abdicate this principal responsibility. Adopting the maxim of degree of regulation commensurate with degree of hazard, along with a product-based trigger, gives greatest assurance to the populace and also brings regulatory resources to bear on the greatest actual threats to public and environmental safety.

LIFE BOX 13.1. ALAN McHUGHEN

Alan McHughen, CE Biotechnology Specialist and Geneticists, University of California Riverside.



Alan McHughen. Even after 40 years, I never tire of reading a DNA sequence. This one is CTTCCTCATGTATATACATGAG. Courtesy of Alan McHughen.

I was a rebel scientist from childhood, eschewing the instructions in my first a chemistry set to "change water to wine" (using phenolphthalein or sodium ferrocyanide) as too lame, preferring instead to blow things up and researching how to use vinegar on the ferrocyanide to release cyanide gas.

Somehow, I survived long enough to harness my enthusiasm for learning how nature worked and set my focus on life, literally and figuratively. Inspirational (and patient) high school teachers taught me how to think critically and how to challenge authority without being obnoxious (a lesson that "took" more often in theory than in practice) and that, although we knew the physical structure and genetic code of DNA, we still didn't understand the essence of life; how, say, DNA controls the sequential unfolding development of a flower, or the fingers on a hand, or how, if DNA uses the same language in all species, nature could derive such magnificent and complex biodiversity.

As a first-year undergrad at Dalhousie University, I was fortunate to land a part-time job with Prof. Gary Hicks, who mentored me in the scientific method and proper experimental protocols when I wasn't washing his Petri dishes or cutting paraffin sections on a microtome. I was not a great academic student, but felt at home in the lab, which allowed me to graduate with decent grades and two peer-reviewed publications. The latter (certainly not the former) catapulted me into a graduate program at Oxford, working with Dr. F.A.L. Clowes researching just how does DNA control the development of a flower.

Doctorate in hand, and still fascinated with this question, I won a research fellowship to continue the work with Prof. Ian Sussex at Yale, where I expanded my technical horizons to apply molecular genetic knowledge to not merely understand nature, but to use her tools to actually do something useful for society. Most of my friends in molecular genetics were drawn to medical applications in the fledgling field of genetic engineering, but I felt my experience working with plants gave me a lift to work in agricultural applications, even though I'd never lived on a farm or even in a farm community—I was the fictitious archetypal city kid who thought chocolate milk came from brown cows.

Leaving Yale, I took a chance. At a faculty job interview with a big agriculture school, I told the department chair, "But I don't know anything about agriculture," and he replied "We'll teach you what you need to know about agriculture if you teach us what we need to know about molecular genetics." It was the early 1980s, and my friends and colleagues knew that genetic engineering would have dramatic impacts in many fields, especially in medicine and agriculture. Surprising even myself, my iconoclastic persona emerged and immediately accepted the job offer.

Twenty years later, my students, colleagues, and I had developed new crop varieties using both genetic engineering and conventional breeding methods. I learned to appreciate how hard farming is, and how important food production was. I also learned that many

people think they understand food and farming, but don't. A hundred years ago, agriculture employed over 40% of the US workforce; today its < 2%, which means 98% of us rely on (usually) unreliable sources on the internet for information.

Now, when I get together with my old friends, I see successful scientists who've contributed to the development of many life-saving or enhancing medical treatments and pharmaceutical products. They've been richly rewarded financially and have rightly earned the respect, if not adulation, of the wider public. Meanwhile, those of us who, 30 years ago, chose to use our skills to help feed the ever-increasing human population with safer, more nutritious foods and feeds, grown with fewer resources and chemical inputs on less land, and to do so in a more sustainable manner, are often seen as pariahs who poison the land and the populace with fearsome Frankenfoods. That the same underlying technology, rDNA, is used in both medical and agricultural applications is lost on the frightened consumers of both insulin and corn, welcoming one and condemning the other. The fact that over 90% of family farmers worldwide who have access to genetically engineered seeds are choosing to grow them, leading to stable increases in safe, affordable food, with no documented harm to any consumer is antithetical to many, especially those who paradoxically claim to "support our family farmers."

I'm sometimes asked, "Do you regret not going into the medical field when you had the chance?" And I instantly and emphatically reply, "No." I could make a lot more money, but I'm still a scientist serving the public, I still love working with nature. And I still want to use my skills, however limited, to best advantage help improve the lot of our fellow humans and preserve what's left of our natural planet. The medical and pharmaceutical fields already have many good scientists, but hungry and malnourished people, whether here or overseas, need more champions working to improve their plight while producing more food in an environmentally sustainable manner. I'll stay where I am, working with my like-minded colleagues, and encourage more young scientists—especially the rebels who aren't afraid of a challenge—to join us.

LIFE BOX 13.2. RAYMOND D. SHILLITO

Raymond D. Shillito, Research and Development Fellow—Seed and Trait Safety at Bayer CropScience LP (USA).



Ray Shillito. Courtesy of Ray Shillito.

How did I end up doing what I am doing now? I followed my instincts, stayed open to possibilities, made mistakes, collaborated with good people, and never stopped learning. My advice is to find something you enjoy doing, as you will usually be good at it. One major thing I learned along the way was, in research, to only try to do one difficult thing at a time. Another is that traditional biochemistry was a great basis for work in this field.

My entry into this field was through tissue culture: a discipline that is still way underestimated by most people. I studied quantitative biochemistry and became interested in obtaining auxotrophic mutants in plants. I was fortunate enough to get a position to do a PhD at the University of Leicester with Professor H. E. Street, a major force in plant tissue culture. Toward the end of my studies, we discussed using insertion mutagenesis with *Agrobacterium* to make mutants and H. E.

suggested the laboratory of Prof. Schilperoort. Sadly, H. E. Street died at Christmas 1977. Lyndsey Withers and Bill Cockburn helped me complete my thesis, and I obtained a grant to study in Schilperoort's laboratory at Leiden in the Netherlands.

When I arrived, Loci Marton had just left, having shown that regenerating *Nicotiana* protoplasts could be transformed by *Agrobacterium*. Thus, I was introduced to Agrobacterium and protoplasts, and to the worlds of gene transfer and molecular biology. When my grant ran out, I used the Dutch I had learned and worked as a postman for 3 months. You have to be able to turn your hand to anything, and it was a great way to meet the real (non-academic) people. I obtained an EMBO grant and moved to Basel, Switzerland, where Ingo Potrykus assembled a team at the Friedrich Miescher Institut to work on transformation and protoplasts; Jurek Paszkowski and Mike Saul completed the core team. Work at the FMI was very collaborative, and I enjoyed working with the groups led by Pat King, and Barbara and Tom Hohn. In 1983, we were able to transform protoplast directly using DNA without any Agrobacterium sequences. This was a major contribution to the field, and we had an excellent experimental system to test other ideas. It led to transforming protoplasts at high efficiency and studies of co-transformation, expression of selectable markers, and of inheritance of introduced genes. My interest in Agrobacterium led to collaboration with Szdena Nicola-Koukolikova and others to investigate the structure of DNA transferred during transformation which was published in Nature. This was a very stimulating time, and I was lucky to experience working with a fun and successful group of people.

After 5 years, Mary-Dell Chilton gave me the opportunity to move to the United States, to Ciba-Geigy's Biotechnology effort. Due mainly to the efforts of Catherine Cramer and Gleta Carswell, we were able to show regeneration of elite maize protoplasts by 1988. Next, my traditional biochemistry training helped in developing a novel selection method for PPT resistance using a pH indicator. We were able to use it to select transformed colonies arising from maize protoplasts and thus regenerate a transformed maize plant. The pH indicator method was then used by Martha Wright and her team to obtain what eventually became the Bt corn event E176. I was involved in filing several patents, which is an interesting pursuit in itself.

When I moved to AgrEvo to build and manage a group to do regulatory studies, I moved from research to development. This opportunity gave me the chance to get to know crop plants in the real world. I carried out the characterization studies that are needed to obtain registration of a biotech crop, and managed development of immunochemical test methods. In 2001, I became involved in the ILSI International Biotechnology Committee, of which I eventually became the chair. This was a great opportunity to learn and to influence how biotechnology was understood by scientists and regulators. We traveled to many countries to hold workshops on the products of the task forces, and I developed a very wide group of contacts. My role increasingly became one of technical expert in the testing field and dealing with outside laboratories and agencies. In this role, I was heavily involved in dealing with withdrawal of StarLinkTM Corn, and with the LibertyLink® rice situation in 2006. I learnt a lot about working under pressure, and now teach a related incident management course. My present work involves establishing and maintaining contacts with a multitude of stimulating people on a daily basis, and planning and presenting at workshops both within and outside of Bayer. I get to use and pass on the skills and knowledge I have learnt along the way, including the quantitative approach to analysis, which means I have in a way come full circle.

Overall, I can say I have enjoyed most of the journey. I retained my links with the tissue culture community for a long time, as this is the basis of modern agricultural biotechnology, but cannot get to the Society for In Vitro Biology (SIVB) meetings any more. I have been fortunate to have the company, advice and mentoring of many people as mentioned earlier, including those who were authorities in their field. I have been lucky enough to follow and grow with a technology

from its inception through to its implementation in agriculture, and have got a great deal of enjoyment out of it. I continue to learn and develop new skills and (I hope) to contribute. In my position as R&D fellow, my job includes taking a wide view and connecting across the silos that often arise in business and academia. I will leave it to others to judge my contribution to the field.

As to where this discipline and plant biotechnology is headed, we are gradually overcoming misinformation and misconceptions. I hope that we can be allowed to use the technology to benefit those who need it most, without over-restrictive regulations, where delivery of critical traits to small subsistence farmers in a safe form of seed which is easily used is a daily occurrence rather than a dream. I was fortunate to be chosen to address the National Research Council Genetically Engineered Crops study in 2014 on this topic. However, this will take time, maybe too much time for some people.

REFERENCES

AAAS (2013): Statement by the AAAS Board of Directors On Labeling of Genetically Modified Foods 2012. Available at: http://www.aaas.org/news/releases/2012/media/AAAS_GM_statement.pdf (accessed on November 4, 2015).

Berg P, Baltimore D, Boyer HW, Cohen SN, Davis RW, Hogness DS, Nathans D, Roblin R, Watson JD, Weissman S, Zinder ND (1974): Letter: potential biohazards of recombinant DNA molecules. *Science* **185**:303.

Clough SJ, Bent AF (1998): Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**: 735–743.

Cohen SN, Chang AC, Boyer HW, Helling RB (1973): Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci USA* **70**:3240–3234.

EFSA (2012a): EFSA panel on genetically modified organisms (GMOs): scientific opinion addressing the safety assessment of plants developed through cisgenesis and intragenesis. *EFSA J* **10**:2561.

EFSA (2012b): Scientific opinion addressing the safety assessment of plants developed using Zinc finger nuclease 3 and other site directed nucleases with similar function. *EFSA J* **10**:2943.

Fraley RT, Rogers SB, Horsch RB (1983a): Use of a chimeric gene to confer antibiotic resistance to plant cells. Advances in Gene Technology: Molecular Genetics of Plants and Animals. Miami Winter Symposia, Vol. **20**, pp. 211–221.

Fraley RT, Rogers SG, Horsch RB, Sanders PR, Flick JS, Adams SP, Bittner ML, Brand LA, Fink CL, Fry JS, Galluppi GR, Goldberg SB, Hoffmann NL, Woo SC (1983b): Expression of bacterial genes in plant cells. *Proc Natl Acad Sci USA* **80**:4803–4807.

Herrera-Estrella L, Depicker A, van Montagu M, Schell J (1983): Expression of chimeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature* **303**:209–213.

Ho, J-Y, Weide R, Mai HM, van Wordragen MF, Lambed KM, Koornneef M, Zabe1 P, Williamson VM. (1992): The root-knot nematode resistance gene (Mi) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. *Plant J* **2**:971–982.

James C (2005): *Global Status of Biotech/GM Crops in 2005*. ISAAA Briefs No. 34-2005. ISAAA: Ithaca, NY.

James C (2013): *Global Status of Commercialized Biotech/GM Crops: 2013*. ISAAA Brief No. 46. ISAAA: Ithaca, NY.

Jia S, Peng Y. (2002): GMO biosafety research in China. *Environ Biosafety Res* **1**:5–8.

Kessler C, Economidis I (2001): *EC Sponsored Research on Safety of Genetically Modified Organisms. A Review of Results*. European Communities, Luxembourg.

MacIlwain C (2003): Chinese agribiotech: against the grain. *Nature* **422**:111–112.

McHughen A (2006): Problems with the Cartagena Protocol. *Asia Pacific Biotech* **10**:684–687.

McHughen A (2007): Fatal flaws in agbiotech regulatory policies. *Nat Biotechnol* **25**:725–727.

McHughen A. (2014): No sustainable human future without biotechnology. In: *Oxford Handbook of Food, Politics, and Society*, Herring, R. (ed). Oxford University Press, Oxford.

Miller J, Bradford K (2010): The regulatory bottleneck for biotech specialty crops. *Nat Biotechnol* **28**:1012–1014.

Murai N, Sutton DW, Murray MG, Slightom JL, Merlo DJ, Reichert NA, Sengupta-Gopalan C, Stock CA, Barker RF, Kemp JD, Hall TC (1983): Phaseolin gene from bean is expressed after transfer to sunflower via tumor-inducing plasmid vectors. *Science* **222**:476–482.

National Research Council (NRC) (1987): *Introduction of Recombinant DNA-engineered Organisms into the Environment, Key Issues*. White paper. National Academies Press, Washington, DC.

Nordlee JA, Taylor SL, Townsend JA, Thomas LA, Bush RK (1996): Identification of a Brazil-nut allergen in transgenic soybeans. *New Engl J Med* **334**:688–692.

NRC (2000): *Genetically Modified Pest Protected Plants: Science and Regulation*. National Academies Press, Washington, DC.

NRC (2002): *Environmental Effects of Transgenic Plants*. National Academies Press, Washington, DC.

NRC (2004): *Safety of Genetically Engineered Foods*. National Academies Press, Washington, DC.

NRC (2010a): *The impact of genetically engineered crops on farm sustainability in the United States*. National Academies Press, Washington, DC.

NRC (2010b): *Strategic Planning for the Florida Citrus Industry: Addressing Citrus Greening*. National Academies Press, Washington, DC.

Organisation for Economic Cooperation and Development (OECD) (1982): *Biotechnology, International Trends and Perspectives*. OECD Publishing, Paris. Available at: http://www.oecd.org/sti/biotech/2097562.pdf (accessed on September 13, 2015).

OECD (1986): *Recombinant DNA Safety Considerations: Safety Considerations for Industrial, Agricultural and Environmental Applications of Organisms Derived by Recombinant DNA Techniques*. OECD Publishing, Paris.

Ryan CD, McHughen A (2014): Tomatoes, potatoes and flax: exploring the cost of lost innovations. In *Handbook on Agriculture, Biotechnology and Development*, Smyth SJ, Phillips PWB Castle D (eds). Edward Elgar, Cheltenham, pp. 841–852.

Schell J, van Montagu M, Holsters M, Zambryski P, Joos H, Inze D, Herrera-Estrella L, Depicker A, de Block M, Caplan A, Dhaese P, Van Haute E, Hernalsteens J-P, de Greve H, Leemans J, Deblaere R, Willmitzer L, Schroder J, Otten L (1983): Ti plasmids as experimental gene vectors for plants. Advances in Gene Technology: Molecular Genetics of Plants and Animals. Miami Winter Symposia, Vol. **20**, pp 191–209.

Smith PG (1944): Embryo culture of a tomato species hybrid. *Proc Am Soc Hort Sci* **44**:413–416.

US Office of Science and Technology Policy (OSTP) (1986): Coordinated framework for regulation of biotechnology. *Federal Register* **51**:23302–23393.

USDA/ERS (2005): The 20th Century Transformation of U.S. Agriculture and Farm Policy. Available at: http://www.ers.usda.gov/media/259572/eib3_1_.pdf (accessed on September 13, 2015).

USDA/ERS (2014): Genetically Engineered Crops in the United States. Available at: http://www.ers.usda.gov/publications/err-economic-research-report/err162.aspx (accessed on September 13, 2015).

Wang G-L, Song WY, Ruan R, Sideris, S, Ronald P (1996): The cloned gene Xa21 confers resistance to multiple Xoo isolates in transgenic rice. *MPMI* **9**:850–855.

Wozniak C, McHughen A (2012): *The Regulation of Agricultural Biotechnology in the United States and Canada*. Springer. Dordrecht, Heidelberg, New York, London.

CHAPTER 14

Field Testing of Transgenic Plants

DETLEF BARTSCH AND ACHIM GATHMANN

Federal Office of Consumer Protection and Food Safety, Berlin, Germany

CHRISTIANE SAEGLITZ

Biotechnology, Bioeconomy, Health Research, Project Management, Forschungszentrum Jülich GmbH, Jülich, Germany

ARTI SINHA

Wayu Health, Gurgaon, India

14.0. CHAPTER SUMMARY AND OBJECTIVES

14.0.1. Summary

When companies or academic labs develop transgenic plants for improved traits, they must assess growth, yield, and trait performance under field conditions. In addition, environmental risk analysis experiments are also commonly performed in the field. A tiered assessment is recognized as being the most appropriate and rigorous approach to assess environmental and economic effects from both scientific and regulatory standpoints. Field design and statistical considerations are described here that are designed to assess the performance of transgenic plants using transgenic maize as an exemplary case study.

14.0.2. Discussion Questions

1. What are the two overarching objectives for field testing of transgenic plants?
2. What two factors are determined for risk assessment?
3. What are some important and appropriate controls for field testing—say, for Bt crops?
4. Give some examples of lower-tier experiments versus upper-tier tests. Why bother with lower-tier tests?
5. Discuss what factors would be needed for the risk assessment of a non-agronomic trait, such as the production of a pharmaceutical. Where would the risk assessor begin, and how would we know when the risk assessment is over—that is, a decision between safe and not safe?
6. Which is more important: that a field test be performed for grain yield or environmental biosafety?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

14.1. INTRODUCTION

Field testing is an important last step in the creation of transgenic plants. Two important and interrelated aspects are discussed here: agronomical performance and biosafety. If a company wants to commercialize a transgenic crop variety (and typically they do), it is important to show that it performs as well as its *parent* or *isogenic* variety in a number of geographic locations and conditions. To be useful as a crop, it cannot have any genetic or phenotypic malformations. So, experiments must be performed to compare growth and yield as well as test for the durability and robustness of the transgenic trait in the field. For example, an insect-resistant plant would be required to adequately express the transgene to prevent herbivory by target insects. Robustness of expression under field conditions is needed to guarantee farmers economic benefits. The second part of field testing, *biosafety*, is more complicated, and it is important to show that a transgenic product is as environmentally benign as its non-transgenic counterpart and commonly used pesticides. This has proved to be crucial for placement on the market since worldwide regulation requires a number of tests to convince regulators to approve transgenic plants as safe (see previous chapter). This chapter focuses on field testing to evaluate the environmental safety (exemplifying any risks for nontarget butterfly caterpillars) and the economic benefit (on yield) posed by genetically modified Bt maize. Thus, we will use this particular case as an example, since it covers much of the ground that is needed for field testing. Some examples, such as those for herbicide resistance, might be simpler, and there are, no doubt, more complicated cases. We use Bt maize since there is a large body of knowledge that has been accumulated over the past 20 or so years and it is a success story of sorts. Since the mid-1990s, genes of *Bacillus thuringiensis* (Bt) that encode butterfly-specific toxins (Cry1Ab, Cry1Ac, Cry1F, and Cry9) were engineered into maize for protection against the European corn borer (*Ostrinia nubilalis*). Bt maize has been hailed as a success, having essentially passed all the regulatory tests in the United States and the European Union (EU), and is grown widely across North America, part of the EU, and in other areas.

14.2. ENVIRONMENTAL RISK ASSESSMENT PROCESS

Environmental risk assessment (ERA) is particularly significant in the context of genetically modified organisms (GMOs). There are some good reasons to be careful when introducing new technologies, in particular when new biopesticides are introduced into the environment. However, it is believed by some concerned people that any (as-yet-unperceived) effects new products have on the environment could be adverse, if not downright "catastrophic." Whatever the starting point is, a scientifically sound ERA is important in the following aspects:

1. Biological properties of the parental unmodified organism (maize in our example)
2. Source of the introduced gene(s), expression, and nature of the gene product (specific "Bt" proteins kill pest caterpillars, but may also affect "lovely" nontarget butterflies)
3. Characteristics of the genetically modified organism, including its performance and impact on the environment, taking into account the information of points (1) and (2)

Environmental risk assessment has a conceptual framework consisting of four basic steps described briefly later: evaluation of need for ERA, problem formulation, controlled experiments and gathering of pertinent information, and finally, risk evaluation. The next steps following would be in the area of managing identified risks, but this is not the purpose of this chapter.

14.2.1. Initial Evaluation (ERA Step 1)

The initial evaluation of need determines whether a risk assessment is required for a specific case. Clearly defining the need as it meets the expectations of the final audience will help in designing the

overall risk assessment and determining how the information will be used and communicated. Common reasons for conducting an ERA include regulatory requirements, scientific inquiry, and scientific responses to public concerns.

14.2.2. Problem Formulation (ERA Step 2)

Once the need for the ERA has been clearly defined, the risk assessment moves forward to the problem formulation phase. In this stage, appropriate risk hypotheses are defined in order to address the scope of the assessment (e.g., whether Bt maize harms lovely nontarget butterflies more than does conventional pest control). Biological aspects of the system, such as the specificity of the mode of action and expression (of the particular genetic trait), the spectrum of Bt activity, and Bt susceptibility of caterpillar as well as relevant exposure profiles are considered while formulating the hypotheses. Other points to consider while identifying potential risks include the intended scale of cultivation (all of the United States or only a few states) as well as other ecological considerations that might affect the environmental impacts (e.g., protected areas with rare or lovely butterflies near cultivation sites).

14.2.3. Controlled Experiments and Gathering of Information (ERA Step 3)

The next step in the ERA involves conducting tests and experiments to gather data pertaining to the study. For example, only a selected group of concerned butterflies can be feasibly studied at one time under laboratory and later field conditions. Hence, species selection must be done very carefully—ensuring that the butterfly species represent both ecologically and economically important taxa. The data collected are used to characterize hazard and exposure.

14.2.4. Risk Evaluation (ERA Step 4)

The overall assessment of the risks is a complicated process. Evaluation of risk would involve the consideration of several perspectives, and can easily go haywire. What is known as the *tiered risk assessment* model was introduced to enable a standardized scientific evaluation of risks internationally. This method consists of several tiers, each consisting of a description of the "problem" at a specific level and the approach to be followed in dealing with it.

14.2.5. Progression through a Tiered Risk Assessment

A *tiered risk assessment* is recognized as being the most appropriate and rigorous approach to assess nontarget effects from both scientific and regulatory standpoints. Both hazards and exposure can be evaluated within different levels or "tiers" that progress from worst-case scenarios framed in highly controlled laboratory environments to more realistic conditions in the field. Lower-tier tests serve to first identify and test potential hazards, and they are conducted in the laboratory to provide high levels of replication and control, which increase the statistical power to test hypotheses. Where no hazards are identified and the transgenic crops are not different from conventional crops, the new product is considered to be as safe as its non-transgenic counterpart. Where potential hazards are detected in these early tier tests, additional information is required. In these cases, higher-tier tests can serve to confirm whether an effect might still be detected at more realistic rates and routes of exposure. Higher-tier studies, including semi-field or field-based tests, offer greater environmental realism, but they may have lower statistical power. Lower statistical power means that the there is a greater likelihood that real effects will not be observed (false negative). One reason for lower power is the high variability of environmental conditions (e.g., climate) that might counteract GM trait-specific effects. Nevertheless, these higher-tier tests are triggered only when early tier studies in the laboratory indicate potential hazards at environmentally relevant levels of exposure. In exceptional

cases, higher-tier studies may be conducted at the initial stage when early tier tests are not possible or meaningful. For example, plant tissue might be used because purified protein is not available for lower-tier work. Higher levels of replication or repetition may be needed to enhance statistical power in certain circumstances. In cases where a potential hazard is detected in a lower-tier test, the tiered approach provides the flexibility to undertake further lower-tier tests in the laboratory to increase the taxonomic breadth (e.g., testing more insect species) or local relevance of test species, thus avoiding the costs and uncertainties of higher-tier testing. Depending on the nature of the effect, one may also progress to higher-tier testing anyway, particularly in cases where there is no previous experience with the crop or transgenic protein under investigation. The various tiered approaches that have been described for *nontarget risk assessment* differ in their specific definitions of individual tiers, but they all follow the same underlying principles. Higher-tier tests usually involve semi-field or field tests and sometimes are conducted when lifecycle (especially reproduction parameters) or *tritrophic* evaluations are warranted. In general, these tests are problematic because of their complexity and high intrinsic uncertainty. Higher-tier tests require expertise and care in experimental design, execution, and data analysis. As a consequence, they are subject to problems of low statistical power, particularly if they are used for "proof of hazard." These tests should therefore be conducted only when they can further reduce uncertainty in the risk assessment, and only when justified by detection of unacceptable risk at the lower tiers of testing. For further reading, see Romeis et al. (2008).

Statistical power has been mentioned several times, and this concept requires clarification. Multiple samples and replicates of experiments are needed for high statistical power, which we can define here as the ability to detect real differences that might exist. Biological systems are highly variable, and statistical tests help researchers test hypotheses, for example, to determine if the differences observed are due to chance variation or result from expression of a transgene. Lower-tiered experiments that can be tightly controlled offer higher capacity to detect real differences than when we layer field effects on higher-tiered experiments. The ground rule is that the more lifelike the experiment, the bigger and more expensive it will be to truly understand natural variability and variability caused by the transgene addition. Larger sample sizes, however, give the experiment higher statistical power.

14.3. AN EXAMPLE RISK ASSESSMENT: THE CASE OF Bt MAIZE

Let us examine the scenario that has garnered the most attention in the risk assessment world: Bt maize pollen exposure. During flowering, maize pollen might land on leaves of host plants (hosts or food for insects) growing in and around maize fields, and these plants might be consumed by caterpillar larvae (ERA step 1). Fields and field margins are important habitats for some butterfly species. As a consequence of the intensification of agricultural practices and the loss of (semi) natural habitat types, field margins have become increasingly important habitats for conserving biodiversity. *Risk* is defined as a function of the adverse effect (*hazard* or *consequence*) and the likelihood of this effect occurring (*exposure*). For butterfly species, the potential hazard is the toxicity of pollen containing Bt protein, and the likelihood of the event is the environmental exposure of caterpillars to the pollen (ERA step 2). In order to quantify hazard, laboratory studies show that monarch butterfly caterpillars that consume Bt maize pollen from the transgenic event Bt 176 had higher mortality, slower development, and lower pupae weights than did those fed non-Bt control pollen (ERA step 3a—hazard characterization). This result caused a great deal of angst, which was accompanied by media and regulatory attention. This case shows that extrapolation of laboratory data to field scenarios can be quite controversial; this case has been among the most (if not *the* most) controversial of all from GM plants. Laboratory tests provide information on toxicity and fitness parameters, but they often represent "worst-case scenarios," which do not reflect field conditions or population processes that operate over farming landscapes. For example, maybe under tier 1 tests

caterpillars were force-fed too much pollen compared with realistic field exposures. Therefore, adverse effects identified in laboratory studies must be verified under field conditions because spatial, temporal, and environmental factors can alter possible adverse effects from, for example, exposure to the Bt protein or temporal overlap between pollen shed and phenology of butterfly caterpillar (ERA step 3b—exposure characterization).

Research in which experimental exposure of insects to Bt protein under field conditions was performed by the authors of this chapter in Germany. In a database survey it, was shown that approximately 7% of the German butterflies (Macrolepidoptera species) occur mainly in farmland areas where maize is grown (for further reading on how this was done, see Schmitz et al. (2003)). The case study summarized in the following addresses some of the issues discussed earlier. In particular, this study attempted to compare the effect(s), if any, of Bt maize on nontarget lepidopteran larvae, with that of conventional insecticides. The suitability and efficacy of the experimental designs and methods used for ERA were also evaluated. It is important that proper comparisons and control treatments be used in ERA experiments to ensure that results are relevant to real agriculture. Since most farmers would spray insecticide instead of simply letting insects eat their entire crop, it is important that ERA for insect-resistant transgenic plants such as Bt maize include comparisons using chemical insecticides, since this is what most farmers use to control damaging insects. There are a few researchers who would like to use, as the main baseline, idyllic conditions that do not exist in much of real agriculture, but these would not be fair, realistic, or useful comparisons.

14.3.1. Effect of Bt Maize Pollen on Nontarget Caterpillars

An experimental maize field in Germany was studied over a 3-year period from 2001 to 2003 (Gathmann et al. 2006). The field was divided into 0.25 ha plots that were surrounded by a strip of conventionally grown corn with a minimum of 4.5 m in width (Fig. 14.1). There were 24 plots in total, on which corn was cultivated in three different ways (or in more precise terminology, *treatments*).

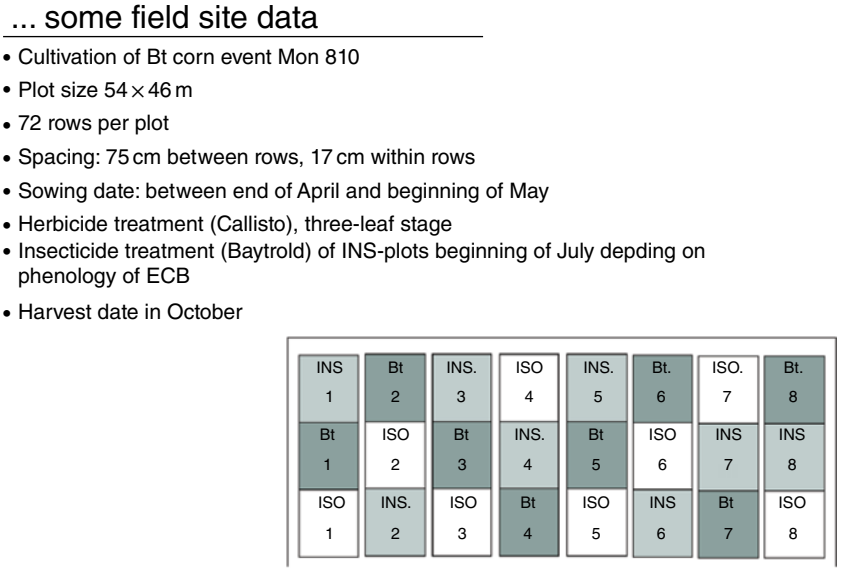


Figure 14.1. Field trial design for testing the environmental impact of bioinsecticidal (Bt) maize pollen or chemical insecticide on nontarget butterflies. The field trial of maize was performed in an area frequently infested with European corn borer (ECB). The trial consisted of eight replications and three treatments (INS = chemical insecticide on a conventional ECB-susceptible variety, ISO = conventional ECB-susceptible variety without any pest ECB control, Bt = GM maize with internal biopesticide protection against ECB).

The maize treatments were used in a randomized pattern to avoid side effects from the surrounding environment. A conventional variety, 'Nobilis' with a similar genetic background but no transgene was used. Recall from Chapter 3 that this is called *near-isogenic*. ISO (O for control) was the control treatment using the near-isogenic plant with no insecticide spray. This treatment provided a baseline for any assessment of effects. In the second treatment the near-isogenic variety was sprayed with the chemical insecticide Baytroid (this treatment is abbreviated INS), which simulated classical pest control. Bt maize, variety 'Novelis' transgenic event MON810 (abbreviated Bt), which synthesized the Cry1Ab protein for insect control, was used in the third treatment. As host plants (weeds) for standardized attraction of butterflies, we used the artificial plantation of goosefoot (*Chenopodium album*) and mustard (*Sinapis alba*) within corn rows (Fig. 14.2). Pollen densities on the host leaves were estimated using a double-sided adhesive tape glued onto microscope slides. Caterpillars were sampled from the plants at the beginning and end of pollen shed. They were carefully replaced back on the plant after identification. Conventional bioinsecticides based on Bt protein have been used for several years in the control of pests, even before the development of Bt transgenic crops.

Studies on Bt bioinsecticide sprays have generally shown negative nontarget effects, that is, predator (parasitoids to the insects) populations were not altered. However, these results cannot be fully extrapolated to transgenic plants producing Bt protein. The microbial Bt products contain Bt protoxins, which are activated in the insect's midgut by proteases (see Chapter 9). Some of the transgenic plants, on the other hand, express partially activated Bt proteins, which could have a potentially different impact on the insect populations. Hence, it can be argued that there is a need to investigate whether the unique delivery system, and the constant exposure of the protein to the insects, has an effect on natural enemies.



Figure 14.2. Weed strip of white mustard (*Sinapis alba*) in a maize field for collection of butterfly larvae on top of weed leaves.

14.3.2. Statistical Analysis and Relevance for Predicting Potential Adverse Effects on Butterflies

Field testing requires careful analysis. For the German field trial on caterpillars, we used a statistical evaluation called the "proof of safety" between Bt maize and the near-isogenic maize variety (ISO). Maize pollen density was estimated to be 52–972 pollen grains/cm^2 on *Chenopodium album* and 100–894 pollen grains/cm^2 in *S. alba*. No significant differences were observed in pollen densities between plant species. Note the wide range of potential exposures. Of the nine butterfly species recovered from the field, only two—*Plutella xylostella* L. and *Pieris rapae* L.—were abundant enough to be considered for statistical analysis (Gathmann et al. 2006). Caterpillars in both of these species are considered to be pests on mustard crops, such as canola, cabbage, and broccoli, but not on maize. Throughout the study period, the numbers of caterpillars (of both *P. xylostella* and *P. rapae*) were lower in plots with insecticide treatment (Fig. 14.3). Pollen density on the plant leaves can be affected by several factors, including relative humidity, growth stage, and distance from maize fields as well as shape and structure of host leaves (e.g., waxy or hairy surfaces). It was observed that more pollen was shed (as inferred from pollen grains/cm^2) from Bt maize in comparison to the conventional maize; however, this could be attributed to the better health of the plants themselves. The Bt plants were observed to be more robust than their isogenic counterparts because they were not damaged by European corn borer as were the ISO plants, which would be expected to lead to the production of more pollen. Hence, no reliable conclusions about the (possible) more adverse effects that they could have on butterfly species could be deduced. No statistically significant detrimental effects of the Bt pollen on the larvae were found (Fig. 14.3). The most important reason for the differences in laboratory results and those from field testing (the latter indicating low overall risk) is the very low level of Bt protein exposure to caterpillar in the field as Bt corn pollen is a much rarer food source under realistic environmental conditions. A less important reason is the temporal overlap between caterpillar development and pollen shed. By the beginning of pollen shed, caterpillars often develop to the final instar stages (Fig. 14.4). Susceptibility to Bt protein is known to decline with caterpillar age, thereby reducing the effect that Bt pollen could have on them.

Similar studies were done on the monarch butterfly to estimate the potential risk under field conditions in the United States. After considering distribution data of the monarch butterfly and their host plants, overlap between pollen shed and development of larvae, and exposure of larvae to Bt

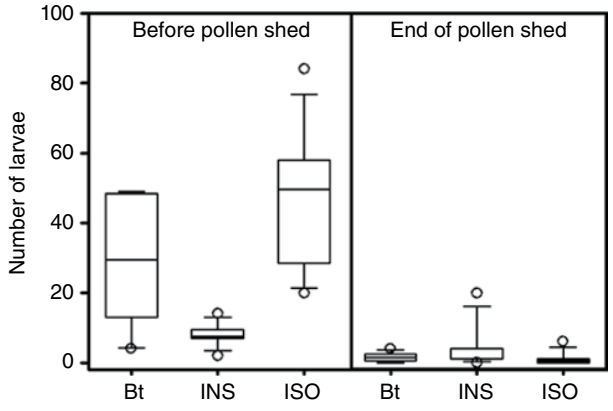


Figure 14.3. Collection of caterpillar larvae on weeds in the maize fields. Presented are the median number of *Plutella xylostella* (diamondback moth) caterpillars on their host plant *Sinapis alba* (white mustard) in the three different treatment Bt maize (Bt) and the near-isogenic variety of maize with (INS) and without (ISO) chemical insecticide application. Larvae were collected before and at the end of pollen shed. Results are pictured as box plots. The horizontal line within each box indicates median of individuals per plot (eight replications for each treatment); the box represents 75% of all values; upper and lower dashes represent 90% of all values. Dots indicate extreme values.



Figure 14.4. Collection of butterfly larvae was done by dislodging them into a cotton cloth tray. Taxonomic knowledge and experience is needed to identify caterpillars to species because of the small size of the larvae.

pollen, risk of Bt pollen to monarch butterflies was determined to be negligible (Sears et al. 2001) (ERA step 4). Again, the amount of Bt maize pollen force-fed to monarch butterfly larvae was much higher than in field exposures, therefore indicating overestimation of risks. It is interesting to note here that the degree of hazard would not change because of exposure. Bt Cry1Ac endotoxin will kill monarch larvae and certain other larvae if there is enough of it. The job of risk assessment is to determine *relevant* exposures and hazards.

14.4. PROOF OF SAFETY VERSUS PROOF OF HAZARD

A "proof of safety" (=equivalence) between Bt maize and the near-isogenic variety was performed using a two-sided (1–2α) confidence test (learn more about this approach by reading Chow and Shao 2002). The percentage change of abundance is easier to interpret than the species-specific absolute difference of arthropods. Therefore, confidence intervals for Bt/ISO ratios were estimated. A ratio >1 for a taxon is equivalent to an x-fold increase in abundance in the treatment; a ratio <1 is equivalent to a decrease in abundance in the treatment down to x%. According to the risk assessment objective, the demonstration of no meaningful population change for selected nontarget species in Bt maize relative to the near-isogenic variety should be shown ("proven"). The population of the nontarget species can be considered as not meaningfully altered if the lower and upper limits of the confidence interval for the abundance ratio are close to and encompass the value 1. Otherwise, the compared treatments cannot be seen as being "equivalent." Abundances can vary in all three treatments; therefore, the confidence intervals for the ratios INS/ISO and Bt/INS were also estimated.

14.5. MODELING THE RISK EFFECTS ON A GREATER SCALE

Mathematical modeling is often a useful tool to estimate potential risks. Since field experiments are done over limited timescales (Perry et al. 2010), it is often useful to model effects over a number of years. The ERA requires estimates of the environmental impact at a landscape or regional scale, and

over longer timescales, such as rotations or decades. These longer experiments are usually not feasible because of limited resources and other factors (Perry et al. 2010). The parameter values chosen for the model were informed by field data where available, and supplemented by educated guesses of experts. Some parameters were generic across species: the proportion of arable fields cropped with maize, the proportion of maize cropped, the average area of fields cropped with maize, and the average width of field margins of fields cropped with maize. Certain parameters were relevant to just weeds: the proportion of the weeds found within arable fields and their margins, the average within-field density of weeds, and the average within-margin density of weeds. Other parameters were relevant to the interaction between caterpillars and pertinent weeds: the degree to which rain washing pollen off weed leaves, caterpillars feeding on the underside of leaves where pollen densities are smaller, and other factors; and the degree to which exposure is reduced due to a lack of temporal coincidence between the susceptible caterpillars concerned and the period over which Bt maize pollen is shed. Two key parameters related to worst-case estimates: the average proportion of susceptible caterpillar that suffered mortality (a) within the field margin of Bt maize fields and (b) within the fields themselves. The results of the modeling were clear: Even for the most sensitive caterpillar species, the (median) estimated mortality over the landscape was only one individual per every 333 larvae. This type of modeling is now routinely applied for ERA of nontarget butterflies, and new updates may be available by using "Perry Bt Model" in search engines.

14.6. PROOF OF BENEFITS: AGRONOMIC PERFORMANCE

When companies or academic labs develop transgenic plants with traits that are shown to be environmentally safe, the plants should still pass a performance test in which growth and yield are assessed under practical field conditions. Again we use transgenic maize as an exemplary case study together with a variety of conventional comparators. Variety registration is a substantial requirement for any new crop brand and varieties in many countries. The evaluation process is governed by independent bodies like the Federal Variety Registration Office in Hanover, Germany. Companies need to send seeds of the new varieties that can be grown on several contracted farms in representative locations in Germany. The comparative approach of yield and performance ensures that farmers get the best available varieties and information. Any new GM maize varieties must be tested in the same way as are non-GM conventional varieties. They will be registered only if their agricultural performance is improved in comparison to standard varieties. Here, we present a representative dataset of field performance of several candidates for variety registration in Germany (Fig. 14.5). In these studies, many plant attributes were tested. Resistance against European corn borer and kernel yield serve as examples, but many traits are considered. The data show that superior yield of GM corn is not always evident in comparison to conventional varieties. Transgenic events need to be integrated into elite variety lines by conventional breeding (see Chapter 3). As we have seen, the genetic background is very important and may also lead to very different performance levels depending on the environmental and climatic conditions at a given site. However, corn borer infestation was dramatically reduced in Bt varieties, leading to various slight yield increases compared to those of three conventional maize varieties. An additional benefit might also stem from the fact that high-corn borer infestations can increase the chance of infections with plant pathogens, such as fungi that produce mycotoxins. Thus, decreasing incidence of pathogens and other indirect factors must be considered as important criteria when evaluating agricultural performance and benefits.

In summary, farmers will buy and cultivate only those elite varieties that fit best their local needs. From the data presented here, it is clear that not only does the level of pest infestation and control determine yield; optimal adaptation to local cultivation conditions, such as soil characteristics, climate, and planting and harvest time, as well as other cultivation practices such as fertilizer and herbicide management, might influence the agronomic benefits for farmers. When field testing is performed, relative crop performance and safety are of utmost importance.

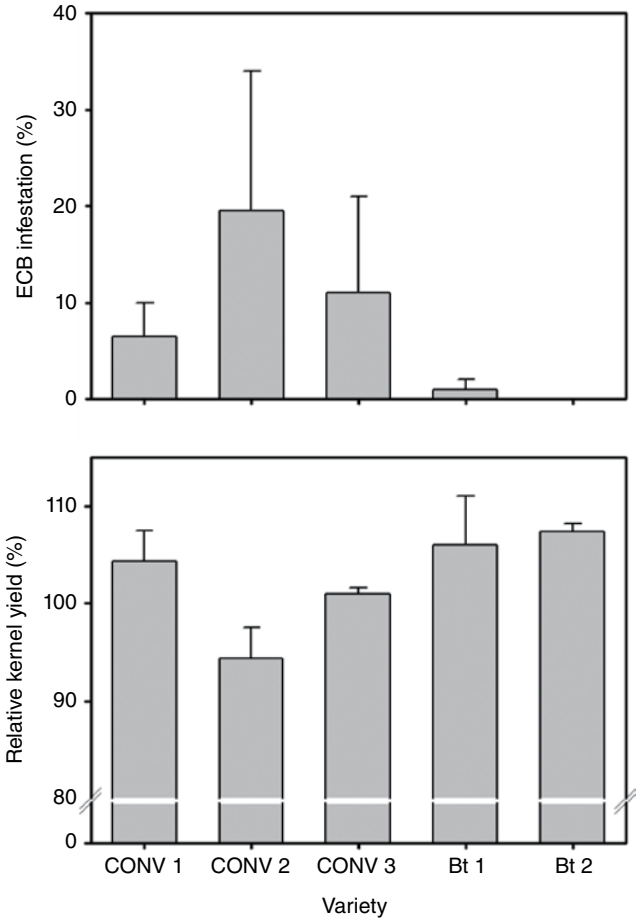


Figure 14.5. Yield performance of three conventional (CONV1–3) and two Bt (Bt1–2) maize variety candidates in a field testing of the Federal German Variety Testing Office (Bundessortenamt-(BSA)). The data on (a) ECB infestation and (b) kernel yield were pooled from fields at three locations in ECB infestation areas spread over Germany. The ECB infestation is specified as percent of infested plants. Relative kernel yield is specified as percentage of yield compared to three standard maize varieties. The bars represent mean values and standard error of the mean.

14.7. CONCLUSIONS

Both biosafety and benefits are important to regulators and consumers. Proof of safety is more important for the regulatory side. Proof of benefits is more important for the economic viability for the company selling the transgenic crop and the farmer who grows it. The example of testing nontarget butterfly species showed the value of field experiments in environmental risk assessment. Laboratory and semi-field studies far overestimated any adverse effect of Bt maize on caterpillars. Field experiments were needed for a comprehensive evaluation of the real environmental effects of Bt maize. As every plant protection practice has an impact on agroecosystems, the overall risk-benefit evaluation needs to compare the impact of both chemical and GM pesticide treatment on nontarget organisms (in this case, butterflies) and the yield performance. Field testing for variety registration demonstrates, on the one hand, the sensitivity of the testing system to agricultural management practices and, on the other hand, the environmental impact of conventional and biotechnological pest management strategies. Both types of environmental and economic studies did not indicate any adverse effect of Bt maize per se. However, not every Bt maize variety had the power to provide necessary (yield) benefits to the farmer.

LIFE BOX 14.1. TONY SHELTON

Tony Shelton, Professor of Entomology, International Professor for Cornell, Associate Director of International Programs, Cornell University.



Tony Shelton. Courtesy of Tony Shelton.

I took a circuitous route in my formal studies as a scientist, but don't regret it. Although I was accepted into a pre-med program, during my first week on campus I transferred into the Great Books of the Western World Program. In this unique program, we had no formal lectures but rather a Socratic dialogue in all our classes. We only read original works, no textbooks. People interpreted the texts, argued their views, and came up with a better understanding based on the discussion. Freshman year, we studied Euclid and the Greek classics and worked our way up to Einstein and Joyce in our senior year. After receiving my BA in classics and philosophy, I went into the biological sciences. Entomology was particularly appealing since it combined my love of ecology, biology, food systems, and the environment. Like many young people at the time, I was tremendously influenced by Rachel Carson's seminal book *Silent Spring*. There had to be a way of producing our food and fiber in a more environmentally responsible manner, and the idea of integrated pest management (IPM) was becoming a buzzword.

IPM focuses on understanding insect–plant interactions within the environment and using host plant resistance and biological control as foundations for managing pests. Over the years, this concept of IPM has become the standard practice. However, we never really had any food plants that were strongly resistant to caterpillars (Lepidoptera) or beetles (Coleoptera) and, in most agricultural systems, biological control couldn't cut it alone, so insecticides continued to play a key role. One interesting insecticide was a bacterium, *B. thuringiensis* (Bt), which could be sprayed on a plant and was strongly promoted by Rachel Carson as an alternative to broad-spectrum insecticides. When caterpillars took a bite of foliage treated with Bt, they were killed by a protein produced by Bt, but this protein did not affect mammals and most other organisms. The problem with spraying Bt was that it was impossible to treat all the surfaces where an insect would feed and growers had to treat often since sunlight quickly broke down Bt. With the advent of genetic engineering beginning in the 1970s, scientists began to see many possibilities for its use in agriculture. One of the first was to insert Bt genes into plants so the plants would produce essentially the same Bt proteins that were in the foliar spray. In 1996, the first genetically engineered Bt plants were commercialized; and by 2014, they were grown on 78.8 million hectares. Finally, we had some plants that were resistant to some caterpillars and beetles! Perhaps, we were on the road forward that Rachel Carson had advocated.

However, the road forward with genetically engineered insect-resistant plants has had a few bumps in it. On the one hand, the adoption of Bt plants has risen incredibly quickly in several countries and has led to dramatic reductions in the use of "harder" insecticides, fewer pesticide poisonings, and improved farmer income. Additionally, the fear that many insects would rapidly become resistant to Bt plants has not materialized even after more than 18 years (this is in stark contrast to nearly all other insecticides). The instances of insects evolving resistance in the field and causing crop failure are few and isolated but show the need to be vigilant to preserve this important technology. The lack of widespread resistance to date may be due to the wisdom of creating many Bt plants with a high enough dose that heterozygosity for resistance would be controlled (it is the heterozygous individuals that drive resistance in a population) and the requirement of having refuges of non-Bt plants so that susceptible

alleles would be maintained in the population. However, for some insect pests high doses and effective refuges have been a challenge and resistance has occurred. Some of our recent work has also shown that natural enemies can play an important role in delaying the evolution of resistance to Bt plants.

Additional fears that Bt genes would spread to wild and weedy relatives and cause environmental havoc and that non-target organisms, especially biological control agents, would be negatively impacted have proven to be unfounded. In fact, Bt plants have advanced the use of biological control because they have reduced the use of broad-spectrum insecticides that are harmful to many biological control agents. However, regulatory issues and acceptance of Bt plants in some countries has been problematic. Bt plants and other products of biotechnology have been called everything from "unnatural and playing God" to "Frankenfoods." If you asked 100 people in the general public who were opposed to genetically engineered plants their reasons for their position, you'd likely get many different answers including questions about long-term food safety issues, corporate control of agriculture, and globalization. Few people would be knowledgeable enough to ask or interpret the technical issues and to analyze the risks and benefits of using this new technology compared to continuing with older technologies for insect management, many of which are far more hazardous. From a scientific standpoint, the environmental and health benefits of Bt plants have been well documented. However, these benefits often get lost in the bigger discussion about biotechnology, and this presents a serious dilemma in a democratic society.

I strongly believe that scientists have an obligation to make their voices heard on important issues such as genetically engineered plants for pest management, but we must do so in a responsible manner. Isn't it our obligation to help inform the public dialogue on these issues? Who else is more qualified to do so?

LIFE BOX 14.2. DETLEF BARTSCH

Detlef Bartsch, Professor, RWTH University (Aachen) and Regulator at Federal Office of Consumer Protection and Food Safety (Berlin).



Detlef Bartsch with GM maize. Courtesy of Detlef Bartsch.

It was in 1977 when I started my involvement in "science" as a 16-year-old political rebel. I got a flysheet on the potential environmental impact of nuclear power stations and immediately felt that I need[ed] to take action based on the dramatic type of selected information, which was provided by—I must say looking backward—a group of concerned citizens. During the following 2 personal years of storm and stress, I joined several environmental and political initiatives trying to protect nature and the environment. Taking part in public discussions I soon became wary of "official experts"—sometimes professors—who explained in scientific terms that… "there is no reason to worry … everything is safe… trust me I am the expert." As my innocent intuition told me the opposite, I decided that I myself should become an environmental expert for protecting the public against unscrupulous industry-paid footmen.

I thought the best discipline for that purpose was biology, and I started my first semester in college in 1980. The "no nukes" time was soon exchanged by the age of "forest decline due to acid rain," which triggered a specialization in the second half of my study toward ecology. I joined ecosystem research in the Institute of Geobotany of the University of Göttingen and took courses in forest ecology, soil sciences and phytopathology. It was in 1985 when I discovered genetic engineering to be a potential threat to the environment. A small group of students started to critically overview the foundation of the research in my university's Genetics Institute, and I joined their discussion in my free time. At that time, there was a strict distinction between molecular biology and ecological sciences. I basically expanded my masters and PhD time (1986–1990) in the field of ecophysiology on the scientific question: Is soil iron availability the driving force for vegetation differentiation into calcifuge and calcicole ecotypes? My answer was probably yes, but there are more multiple cause–effect relationships. This draws my attention to a personal experience: I am sure that to a large extent in public perception "the example seems to be everything" in ecology. That's probably the reason why one can find for every real or fictional environmental concern support by some ecological data. It is still the great challenge in ecological sciences to find generalizations and rules, which is difficult as there are so many influencing factors.

Anyhow, I continued to be (politically) interested in the environmental consequences of genetically modified organisms (GMOs). It was a lucky random event that paved my way toward environmental biosafety research: I went in the middle of my PhD work in 1988 for a 6-week internship to the German Parliament where I accompanied a parliamentarian engaged in environmental politics. One small task I got was to evaluate a new draft legislative act on GMOs. I made a phone call to the leading German environmental expert in the National GMO Biosafety Committee: Prof. Herbert Sukopp, who was an expert in exotic plant species ecology. A 20-minute chat with him resulted in my being offered a postdoc position in his lab at the Berlin Technical University 2 years later, in 1990. My task was to organize a conference and ecology expert database for the interdisciplinary assessment of GMOs. During the next 15 month, I gained experience and made the right contacts to become involved in the first biosafety research projects with sugar beet in cooperation with plant breeding industry in field experiments starting in 1993. This was the first time that a GM crop was released into the environment in Germany, and I was the first ecologist to study competitiveness and GMO out-pollination with wild-type plant relatives. My next 10 years were characterized by intensive experimental studies, teaching and field trips with students, and highly polarized public discussions with concerned citizens.

In a world of simply black-and-white views, I suddenly was pushed to the "pro-GMO" side as some people were not able to see why it was important to collect scientific data that enables science-based decisions. What a change! I became the opposed official expert myself who is mistrusted by gut-feeling-driven opponents of a new technology. This was one of the most pervasive experiences in my life. It was now my problem to tell people that the truth in the GMO world is colorful and not a black-and-white story. Anyhow, I tried my best both in communicating my research as well as improving my scientific knowledge. A great time in this respect was my sabbatical study on the origin of Californian wild sea beet (*Beta vulgaris* ssp. maritima) in the lab of Prof. Norm Ellstrand at UC Riverside in 1998.

Back in Germany, the environmental impact of Bt corn became a new object of interest. I spent three more fruitful years with my scientific mentor Prof. Ingolf Schuphan at the Aachen University of Technology. The German university system has a narrow window of opportunity to obtain a professorship. Even though I was near the final cut, I had no luck in the end to get a full professorship position.

But as luck would have it, I left in 2002 for a full-time regulator job in a Federal German Agency. My job is highly inspiring since I combine scientific background information, political implications, and cost-benefit considerations into regulatory decisions. I've learned during my scientific career that good decisions are mostly those that are taken based on knowledge and not on uncertainty. Now I am an expert working for the

German government and as independent expert for the European Food Safety Authority and sit in front of skeptic young rebels who want to save the world against evil techniques, but I think plant biotechnology could potentially offer more advantages than disadvantages for better and more environment-friendly agriculture. Plant biotechnology is based on my on scientific experience really not black-and-white, but is as colorful as life. I hope to be an honest mediator and decision maker.

REFERENCES

Chow SC, Shao J (2002): A note on statistical methods for assessing therapeutic equivalence. *Contr Clin Trials* **23**:515–520.

Gathmann A, Wirooks L, Hothorn LA, Bartsch D, Schuphan I (2006): Impact of Bt maize pollen (M0N810) on lepidopteran larvae living on accompanying weeds. *Mol Ecol* **15**:2677–2685.

Perry JN, Devos Y, Arpaia A, Bartsch D, Gathmann A, Hails RS, Kiss J, Lheureux K, Manachini B, Mestdagh S, Neemann G, Ortego F, Schiemann J, Sweet JB (2010): A mathematical model of exposure of non-target Lepidoptera to Bt-maize pollen expressing Cry1Ab within Europe. *Proc Royal Soc B Biol Sci* **277**: 1417–1425.

Romeis J, Bartsch D, BiglerF, Candolfi MP, Gielkens MMC, Hartley SE, Hellmich RL, Huesing JE, Jepson PC, Layton R, Quemada H, Raybould A, Rose RI, Schiemann J, Sears MK, Shelton AM, Sweet J, Vaituzis Z, Wolt JD (2008): Non-target arthropod risk assessment of transgenic insecticidal crops. *Nat Biotechnol* **26**:203–208.

Schmitz G, Pretscher P, Bartsch D (2003): Selection of relevant non-target herbivores for monitoring the environmental effects of Bt maize pollen. *Environ Biosafety Res* **2**:117–132.

Sears MK, Hellmich RL, Stanley-Horn DE, Oberhauser KS, Pleasants JM, Mattila HR, Siegfried BD, Dively GP (2001): Impact of Bt corn pollen on monarch butterfly populations: a risk assessment. *Proc Natl Acad Sci USA* **98**:11937–11942.

CHAPTER 15

Intellectual Property in Agricultural Biotechnology: Strategies for Open Access

MONICA ALANDETE-SAEZ, CECILIA CHI-HAM, SARA BOETTIGER AND ALAN B. BENNETT

Public Intellectual Property Resource for Agriculture, Department of Plant Sciences, University of California, Davis, California

GREGORY GRAFF

Department of Agricultural & Resource Economics, Colorado State University, Fort Collins, Colorado

15.0. CHAPTER SUMMARY AND OBJECTIVES

15.0.1. Summary

This chapter seeks to explain intellectual property considerations in agriculture and especially the role of patent protection for transgenic plants. The law dealing with patents and the landscape of protecting intellectual property while maintaining public good and enabling humanitarian causes are not simple issues. Nonetheless, the goal of this chapter is to provide the reader some potential scenarios of how "open-source" sharing models could facilitate leveraging technologies for the greater good.

15.0.2. Discussion Questions

1. What is intellectual property, and how can it be protected?
2. What is a patent?
3. What conditions must be met for an invention to be patentable?
4. Why can living things and their components be patented?
5. What is meant by the "tragedy of the commons?"
6. What is meant by the "tragedy of the anti-commons?"
7. What is freedom to operate (FTO)? Why is it important?
8. Is open-source or open-access plant biotechnology feasible?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

15.1. INTELLECTUAL PROPERTY AND AGRICULTURAL BIOTECHNOLOGY

Scientific advances in many fields have been treated historically as public goods, and this has been particularly true in agriculture. Universities and other public-sector institutions were the leaders in developing improved crop varieties that were transferred to farms through cooperative extension services in the US or equivalent organizations internationally (Conway and Toenniessen 1999). This model, however, has changed rapidly in the last few decades, primarily because of greater utilization of formal intellectual property (IP) protection of agricultural technologies and plant varieties by the public sector, as well as the development of a research-intensive private sector that now make major contributions in enhancing the productivity of US agriculture (Kowalski et al. 2002). In particular, the expanded use of formal IP rights for agricultural biotechnology-based products can be understood by considering the significant amount of time and financial resources needed to develop a new transgenic crop and the high costs of obtaining regulatory approval to market such a crop. In the face of these costs, the time-limited exclusivity provided by patents allows the investor an opportunity to recoup the costs of research and development. Indeed, it is very likely that the agricultural biotechnology industry would not have developed in the absence of a strong framework for IP protection.

The growth in patents related to agricultural biotechnology can be seen in Figure 15.1. These data indicate a strong growth in the issuance of patents by the US Patent and Trademark Office, and similar trends are also apparent in patent applications internationally, suggesting that this is a global trend. The scope of inventions represented by the data in Figure 15.1 is quite broad but can be conceptually divided into two main categories: (1) those that cover research tools or "enabling technologies" that are required to produce transgenic plants or to discover new gene functions and (2) those that cover "trait technologies" that confer specific attributes to genetically modified plants.

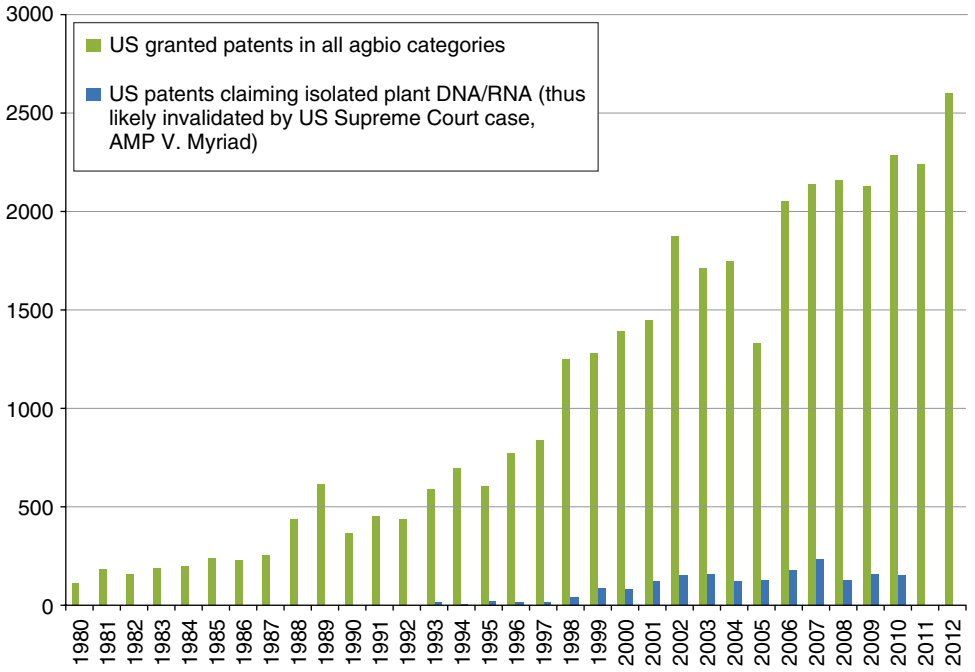


Figure 15.1. Annual trends in plant biotechnology based on patents granted by the US Patent and Trademark office between 1980 and 2012. Analysis includes US patents granted and corresponding to the international patent class categories considered relevant in agricultural biotechnology (gray bars, Data Source: Thomson Innovation 2013; http://info.thomsoninnovation.com). Included are US patents claiming isolated DNA/RNA sequences that are likely to become invalidated by the US Supreme Court. (*Source*: Adapted from Graff et al. (2013).)

This distinction is important because all researchers and research institutions need access to the fundamental tools of agricultural biotechnology if the greatest benefits of the technologies are to be realized, whereas exclusive access to specific trait technologies is an effective means of ensuring that the new crop varieties expressing these attributes are developed. As a consequence, there is a delicate balance in the overall innovation framework between exclusive access to certain technologies while at the same time ensuring broad access to other technologies. A very similar situation was addressed in the early 1980s when Stanford University and the University of California, San Francisco, patented the basic methods of recombinant DNA manipulations (Cohen and Boyer 1980). This patent, covering the fundamental tool of modern biotechnology, was non-exclusively licensed under reasonable financial terms, a strategy that resulted in licenses to 468 companies (Feldman 2005). Broad innovation was encouraged through access to the key enabling technologies, and many companies became successful by using these tools to develop proprietary products based on other patented technologies that they exclusively held. Ultimately, the licensing strategy enabled $35 billion in worldwide product sales and brought in $255 million in licensing revenues (Feldman 2005). For agriculture, it might be helpful to envisage enabling technologies as being upstream, and perhaps necessary, to develop the downstream trait technologies.

In agricultural biotechnology today, the innovation system needs a balance of both exclusive and nonexclusive access to patented technologies to effectively support new crop development and to provide both commercial growers and subsistence farmers with the best genetic technology possible for their crops. The ownership of critical intellectual property and the rights to practice or use certain technologies is becoming a major issue confronting researchers in this area. Even purely fundamental academic research is not protected by an "experimental use" exception from patent infringement and may become increasingly entangled in issues involving access to IP rights (Eisenberg 2003). While the importance of intellectual property in agriculture is becoming better recognized in both the public and private sectors, many researchers, business people, and R&D decision makers and policy makers are still relatively uninformed about how to find, understand, and utilize IP information, including published patents and patent applications. In this chapter, we will provide an overview of the major issues and what a research scientist needs to be aware of when navigating the IP landscape of agricultural biotechnology.

15.1.1. What is Intellectual Property?

Intellectual property is a legally created form of property that applies to ideas or the "products of the mind," which gives the owner a set of rights that are comparable to tangible property rights. The concept of intellectual property was insightfully addressed by Thomas Jefferson when he said, *If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself.…Inventions then cannot, in nature, be a subject of property. (However) society may give an exclusive right to the profits arising from them, as an encouragement to men to pursue ideas which may produce utility* (Jefferson 1987). Jefferson's concept of society providing a legal mechanisms for inventors to have an exclusive right to profits from their ideas was subsequently integrated into the US Constitution, Article I, section 8 which states that *The Congress shall have power … to promote the progress of science and useful arts by securing for limited times to authors and inventors the exclusive right to their respective writing and discoveries*. This forms the basis for IP rights and has become the cornerstone of the innovation process in the United States and, more recently, in many other countries throughout the world.

15.1.2. What is a Patent?

There are several forms of intellectual property, including plant and utility patents, copyright, trademarks, and trade secrets. In agricultural biotechnology, the dominant forms of intellectual

property are patents and these are the primary focus of this chapter. Patents provide just what the constitution promised—the right to exclude others from using your invention. Importantly, this right is conferred by a national government for a specified time period, usually 20 years. So, in general, a patent provides an intellectual property right that is geographically limited to the specific countries in which patent protection is obtained and it is time limited by the term of the patent. This is a significant way in which patent-protected intellectual property differs from tangible or real property where ownership is usually not limited by either geography or time. These differences between intellectual property and tangible property often have an impact in biological research since research materials (vectors, genes, cell lines, etc.) are usually obtained under the terms of a material transfer agreement (MTA) that likely contains provisions on how the material is used. Because the MTA governs the transfer of tangible or real property, the terms of the agreement typically do not contain geographical or temporal limitations and, as a result, the restrictions imposed by MTAs can become particularly problematic.

The monopoly that a patent provides to an inventor is a very powerful economic right and, as a consequence, the invention must meet a standard of novelty, nonobviousness, and utility. That is, the invention must be original and not previously known, the intervention must not be an obvious extension of previously known information, and the invention must have some useful purpose. The standard of novelty has an important implication for researchers since the primary means of scientific communication is through broad publication, which if done carelessly, can destroy the patentability of an invention. The section of US patent law relevant to novelty says that a patent application can be rejected on lack of novelty grounds if *the invention was… patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for patent, or… more than one year prior to the date of the application for patent in the United States*. In most other countries, the 1-year grace period provided in the United States does not exist and public disclosure of an invention immediately bars patentability in those countries. This is especially important since March 16, 2013, when the Leahy–Smith America Invents Act (AIA) was implemented under the US patent law. The AIA transformed the US patent system from a first-to-invent system to a first-to-file system involving changes that make US patent law more consistent with the European Union and much of the world, thus harmonizing the patent systems of different nations to yield greater consistency and predictability in obtaining and enforcing patent rights. Contrary to the former US patent law in which only the domestic activities were relevant in establishing prior art, the new first-to-file system places an emphasis on public disclosure that bars patentability regardless of whether disclosure occurred inside or outside of the United States. Overall, the new first-to-file regime provides an incentive to file a patent application as early as possible since it has significantly increased what it is considered under US law as "prior art," and therefore making the novelty condition a bigger challenge to meet. In addition to the timing of public disclosures, a researcher also needs to consider the meaning of the words *printed* and *publication*. For example, is a document or a slide presentation posted on the Internet considered *printed*, such that the document bars future patentability? Any disclosure of a potentially patentable idea should be made thoughtfully and/or in consultation with an attorney or technology transfer office. In some cases, a clear public disclosure can be purposely designed to bar patentability in order to ensure that an invention remains in the public domain and available for everyone to use without restriction (Boettiger and Chi-Ham 2007).

The patenting of plant and animal genes has been particularly controversial, and critics have argued that genes are not patentable because they exist in nature. In past rulings, the US Patent and Trademark Office (USPTO) concluded that an isolated and purified DNA molecule that has the same sequence as a naturally occurring gene is eligible for patent protection because it does not occur in its isolated form in nature. However, on June 13, 2013, the Supreme Court ruled in the closely watched Myriad gene patents case that the isolation of natural DNA—in this case human genes—was unpatentable on the basis that products of nature cannot be patented. The rationale of the ruling was that the location and order of the nucleotides existed in nature before Myriad discovered them. This ruling has major implications because it will no longer be possible in the

United States to patent or exclude others from using a genomic DNA sequence. However, the Supreme Court also ruled that a cDNA sequence is patent-eligible because it is an exon-only DNA molecule that does not occur in nature. More importantly, the decision does not eliminate the ability to patent genetic diagnostic tests that utilize or rely upon genomic DNA sequences. In addition, the USPTO modified and adopted a higher standard of "utility" in its guidelines for evaluating gene patents, requiring that the applicant demonstrate that the "utility is specific, substantial, and credible" (http://www.uspto.gov/web/offices/com/sol/notices/utilexmguide.pdf). Despite this specific utility requirement, there are a number of patent applications that claim the sequences of hundreds of genes for which the utility is only broadly defined.

For example, US patent application 20070022495 defines the utility of several hundred claimed genes as conferring an "improved trait relative to a control plant" and "the improved trait is selected from the group consisting of larger size, larger seeds, greater yield, darker green color, increased rate of photosynthesis, more tolerance to osmotic stress, more drought tolerance, more heat tolerance, more salt tolerance, more cold tolerance, more tolerance to low nitrogen, early flowering, delayed flowering, more resistance to disease, more seed protein, and more seed oil relative to the control plant." In this specific case, the USPTO limited considerably the scope of the claims of this broad application in the US patent 7,858,848 issued on December 28, 2010. The granted patent claims a transgenic plant expressing a specific polynucleotide conferring increase resistant to a fungal pathogen. It is noteworthy to mention that this patent family contains 15 granted US patents and 382 filings spreading over 10 countries. Time will tell how patent offices and patent examiners treat these broad patent application and patents, and ultimately whether such broad gene patents are enforceable. Furthermore, based on the US Supreme Court Myriad case ruling, there are a number of US granted patents that include claims to isolated nucleic acids which will likely be invalidated (Fig. 15.1).

15.2. THE RELATIONSHIP BETWEEN INTELLECTUAL PROPERTY AND AGRICULTURAL RESEARCH

The impact of public sector research in agriculture has been very significant. In the United States, this dates back to the establishment of the Land Grant College system of universities which have led to the development of improved crop varieties that were transferred to farms and to the agricultural industry through cooperative extension services in the United States Internationally, the system of crop research centers sponsored by the Consultative Group on International Agricultural Research (CGIAR) has a similarly large impact in developing new crop cultivars and agronomic practices that were delivered as a public good to support global food production. This model has been slowly changing and the rate of change is now accelerating. At the core of this change is the increasing role of IP protection over agricultural inventions, as well as the development of a research-intensive private sector in agricultural biotechnology. Thus, both US and global agricultural systems are experiencing a change from research results being developed primarily in the public sector and the resulting technologies delivered for free as a public good to a system that is increasingly dominated by private companies who protect and treat results as a private asset. This has been accomplished through a much more intensive use of the patent system to protect agricultural innovations than was previously the case. The trend in patents awarded related to plant biotechnologies between 1980 and 2012 clearly illustrates the overall increase in patent activity in this sector (Fig. 15.1).

In the past 25 years, other fundamental changes in the nature and ownership of innovations in basic and applied agricultural research have complicated the mission of public research institutions. The primary change was the passage of the Bayh-Dole Act which encouraged US universities to patent their innovations and license them to private-sector companies in order to encourage their commercial use. Since that time, patenting by public research institutions and universities and the development of formal technology transfer mechanisms have accelerated. While public sector

institutions contribute to only about 2.7% of patents overall, their contribution to agricultural biotechnology patents is nearly an order of magnitude greater—contributing approximately 24% of all patents (Graff et al. 2003). While this trend has contributed to many positive economic outcomes, these new policies have also created challenges for public research institutions and universities in supporting broad innovation, particularly for agricultural applications that address small markets such a specialty crops or that support humanitarian, rather than commercial, purposes.

15.3. PATENTING PLANT BIOTECHNOLOGY: HAS AN ANTI-COMMONS DEVELOPED?

The proliferation of patents in biotechnology led to the development of a metaphor to explain why people overuse or under use resources. The "tragedy of the commons" was a term coined by Garrett Hardin to explain why people overused shared resources, such as common pastures, because they have no incentive to conserve or extend to the life of the resource (Hardin 1968). By analogy, Heller and Eisenberg (1998), described the "tragedy of the anti-commons" which, as the result of a proliferation and fragmentation of IP ownership across multiple owners, prevents any single institution or company from assembling all of the necessary rights to produce a product, resulting in the underuse (or nonuse) of resources. Interestingly, whereas patents and IP, generally, are intended to encourage investment in research and development, the development of an anti-commons has the opposite effect of blocking innovation. Although this concept of the anti-commons was initially described in relation to biomedical research, it also has direct relevance to agricultural biotechnology.

A prominent example of the complexity resulting from fragmented technology ownership and the potential for anti-commons to arise were exemplified in the development of β-carotene-enriched rice by public-sector researchers who used at least 40 patented or proprietary methods and materials belonging to a dozen or more different owners in the gene transfer process (Kryder et al. 2000). Some examples of the types of patented technologies that are required for developing a genetically engineered crop include transformation methods, marker genes, trait genes, regulatory elements, and other molecular biology-based inventions.

15.3.1. Transformation Methods

The development of transgenic varieties typically relies on either *Agrobacterium*-mediated or biolistic-mediated gene transfer methods (Herrera-Estrella et al. 1983; Klein et al. 1987). Fundamental methods related to both processes of gene transfer into plant cells were invented in the public sector (Sanford et al. 1991; Barton et al. 2000), but key patents for *Agrobacterium*-mediated transformation were licensed exclusively to Ciba-Geigy (now Syngenta), and the biolistic technology was licensed exclusively to DuPont for most fields of use. In addition, both private-and public-sector R&D organizations have patented a number of fundamental transformation methods, as well as improvements including vectors, species-specific protocols and novel strategies to remove selectable markers and other "foreign" DNA from the plant to be commercialized (Fraley et al. 1991; Hoekema et al. 1992; Hamilton 1998; Yoder and Lassner 1998; Rogers and Fraley 2001; Pray and Naseem 2005). As a result of a variety of transactions, fundamental methods of gene transfer to plant cells were invented either by private-sector companies or by public-sector companies, but then licensed exclusively to private companies and represent a key technology area where patents have the potential to block new innovations. An example is the recently granted US 8,273,954 (Rogers et al. 2012) patent covering a broad *Agrobacterium*-mediated transformation method for dicotyledonous plants, which is owned by Monsanto and has an anticipated expiration date in 2029. The effective filing day of the initial patent application was in January 17, 1983; however, the pending application was not granted for nearly 30 years because of an interference case in which two applicants had claimed the same invention. Finally, in September 2012, Monsanto was granted the patent,

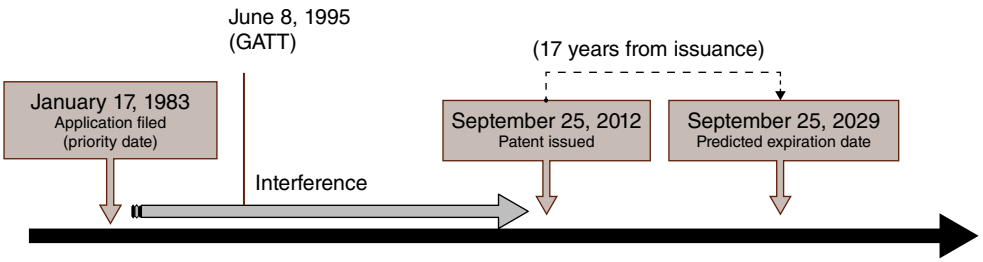


Figure 15.2. Prosecution of Monsanto's *Agrobacterium*-mediated transformation of dicot plants US patent 8,273,954. While the patent application was filed in 1983, prosecution was delayed because of an interference case. The patent was awarded in September 25, 2012; however because the application date was prior to June 8, 1995, the expiration date is calculated at 17 years from the issue date. In June 8, 1995, the US patent law passed a major change to harmonize US law with other countries' law, the General Agreement on Tariffs and Trade (GATT) Treaty. Prior to June 8, 1995, a US patent had a term of protection lasting 17 years from the date of issuance. Under GATT changes, an unexpired issue patent or a patent application pending on June 8, 1995, has a term of protection the longer of 17 years from the date of issuance of the patent or 20 years from the filing date of the patent application, which is the case, is for the US 8,273,954 patent. Any patent application filed on or after June 8, 1995, has a term of protection that begins on the date of the grant of the patent and ends on a date 20 years after the filing date of the patent application.

which has a term of 17 years from the issuance date under US patent law, based on the fact that the original patent application was filed before June 8, 1995, the pre GATT-TRIPS filing date (Fig. 15.2). The scope of Monsanto's patent is a method for transforming dicotyledonous plant cells with any coding sequence genes, including selectable markers, using *Agrobacterium* and a Ti plasmid with *Agrobacterium* T-DNA borders. Monsanto Company announced that it will provide a royalty-free research license to the academic community and other nonprofit research institutions for the *Agrobacterium*-mediated transformation method described in their newly issued patent US 8,273,954. The company also announced that it will continue to make commercial licenses available.

There has been a growing interest in non-*Agrobacterium*, bacteria-mediated transformation. One of the first viable alternative plant transformation methods to *Agrobacterium*-mediated transformation and biolistics was the Transbacter™ suite, developed by the Center for Application of Molecular Biology for International Agriculture (CAMBIA). Transbacter was introduced in 2005 by CAMBIA as a new plant transformation technology using non-*Agrobacterium* bacteria. CAMBIA filed a number of patents; however, the portfolio has been abandoned and the technology is now in the public domain. Recently, a research group based at Teagasc Crop Research Center (The Irish Agriculture and Food Development Authority) discovered an alternative bacterial plant transformation method to the classical *Agrobacterium*-mediated transformation using *Ensifer adhaerens*, a gram-negative soil bacterium capable of a horizontal gene transfer that is genetically distinct from *Agrobacterium*. *Ensifer*-mediated transformation (EMT) exhibits transformation efficiencies that are comparable with *Agrobacterium*-mediated transformation, and more importantly, can be directly substituted into existing *Agrobacterium* protocols without further optimization and may be used with both dicot and monocot plants. The Tegasc EMT patent portfolio currently includes an international patent application WO 2011076933 and three pending patent applications, US 2013,0078,706, CA2784550, and IN201201460 in the United States Canada, and India, respectively. These public-sector inventions may represent alternatives for plant transformation.

15.3.2. Selectable Markers

The most commonly utilized plant selectable marker genes include the *nptII* and *hpt* genes that confer antibiotic resistance as a basis to select for cell transformation (Miki and McHugh 2004)—see Chapter 10. Several other selectable markers conferring herbicide resistance or positive selection

based on novel carbon utilization pathways provide important alternatives to the antibiotic-based selection strategies (Roa-Rodriguez and Nottenburg 2003). Broad patents cover all of these selectable markers (Bojsen et al. 1998; Santerre and Rao 1988; Rogers and Fraley 1993, 2001). Selection strategies appear not to have been the topic of public-sector research programs, and there are just a few examples of either public domain or public-sector-patented selectable markers for use in plant transformation (Dirk et al. 2001, 2002; Mentewab and Stewart 2005). While there is potential to invent new selectable markers for plant transformation, at this point this represents another key enabling technology where patents have the potential to block new innovations.

15.3.3. Promoters

15.3.3.1. Constitutive Promoters. Genetic regulatory elements are required to drive the expression of selectable marker genes and of specific transgenes. Selectable marker genes are typically driven by high-level constitutive promoters with the most common constructs utilizing the CaMV 35S promoter derived from a viral genome and owned by Monsanto (Odell et al. 1985). There are many alternative promoters that confer constitutive gene expression that were developed in public-sector organizations and that are either in the public domain or can be licensed for nominal fees. These alternatives include monocot and dicot actin promoters (McElroy et al. 1990; An et al. 1996; Huang et al. 1997; McElroy and Wu 2002), a FMV 34S promoter (Comai et al. 2000), mannopine/octopine synthase (Gelvin et al. 1999), or FMV and PCLVS FLt promoters (Maiti and Shephard 1998, 1999). The FMV 34S has been used to drive constitutive gene expression and reported to be essentially equivalent to the CaMV 35S promoter (van der Fits and Memelink 1997; Romano et al. 1993), but has not been widely distributed to the public-sector research community. Each of these promoters provides a strategy for driving constitutive transgene expression using public-sector-derived or public domain components.

15.3.3.2. Tissue- or Developmental-Stage-Specific Promoters. Although many genes can be expressed under the control of constitutive promoters, targeting of expression to plant organs or tissues is typically desirable to minimize nonspecific effects of the introduced gene. For example, seed-specific promoters (Blechl et al. 1999; Harada et al. 2001) have been patented with claims directed toward their use to drive expression of heterologous genes in developing seeds. Public-sector institutions have also patented a relatively large number of tissue- and/or developmental-stage-specific promoters. Examples include the root-specific CaMV 35S fragment A promoter (Benfey and Chua 1992), a root cortex-specific promoter (Conkling et al. 1998), the Pyk10 root-specific promoter (Grundler et al. 2001), an epidermal cell-specific Blec promoter (Dobres and Mandaci 1998), and a vascular tissue-specific promoter RTBV (Beachy and Bhattacharyya 1998). In addition, there exists a large number of tissue and developmental-stage-specific promoters that have been characterized and placed in the public domain through publication. A wide range of constitutive and regulated promoters have been tabulated in a promoter database that includes information on expression characteristics as well as their IP status (database hosted by PIPRA: www.pipra.org).

15.3.4. Subcellular Localization

In addition to specificity in tissue-level transgene expression, it is also often important to direct the targeting of the new protein to a specific subcellular location. For example, because β-carotene is produced in the plastids, the development of β-carotene-enriched rice utilized a transit peptide derived from the small subunit of Rubisco to target proteins to this subcellular compartment (Ye et al. 2000). This and other transit peptides have been the topic of intense study, and several companies have patented their use to direct proteins into plastids (Herrera-Estrella et al. 2000; Dehesh 2002). However, several early publications from public-sector research organizations described alternative transit peptides that were not subsequently patented and thus should be accessible in the public domain (Smeekens et al. 1986). Because transit peptides do not have a high degree of sequence

similarity, it is likely that additional transit peptides will not be dominated by existing patent claims and alternative sources of functional transit peptides could be developed from public domain information or from public-sector laboratories. Targeting to other subcellular locations has been the topic of intense research in both the public and private sectors and there are many examples of public-sector research describing unpatented sequences targeting proteins to a variety of subcellular sites including the cell wall, vacuole, plastids, and peroxisomes (Bednarek et al. 1989; Tague et al. 1989; Volokita 1991; Hayashi et al. 1996; Kato et al. 1996; Komarnytsky et al. 2000).

15.3.5. The Importance of Combining IP-Protected Components in Transgenic Crops

Developing a new genetically engineered crop requires the assemblage of a number of patented technologies through in-licensing or, potentially, by a series of strategic mergers and acquisitions. Several companies have effectively done this and have used platforms of proprietary technologies to develop new varieties of major crops. However, work on crops of less commercial interest has progressed slowly with few of the benefits of biotechnology having been realized in specialty crops (Clark et al. 2004). Based on the requirement for assembling a large number of patented technologies to produce genetically engineered crop and the fragmentation of IP ownership, it appears that the preconditions for the development of an anti-commons exist in this technology sector. In addition, the observed slowdown in the development of new agricultural biotechnology products may be, at least in part, an effect of such an IP anti-commons (Graff et al. 2009).

15.4. WHAT IS FREEDOM TO OPERATE (FTO)?

15.4.1. The Importance of FTO

Navigating the complex IP landscape of a research project in agricultural biotechnology, especially for desired commercialization, requires some analytical tools and specialized analytical capabilities (Fenton et al. 2007). The analysis requires both legal and scientific knowledge and access to both patent and literature databases and typically takes the form of what is known as a FTO opinion. The FTO opinion is a legal assessment about whether research project or the development of a new commercial product can proceed with a low, or tolerable, likelihood that it will not infringe existing patents or other types of IP rights. It is important to note that the FTO determination is not absolute but reflects an evaluation of risk since there is typically some uncertainty around the interpretation of patent claims as well as uncertainty as to whether new IP may issue or be discovered at a later date. The FTO opinion may lead to a range of options: identifying in-licensing targets, considering the substitution of technologies, deciding to ignore the potential infringement, investing in work-around technologies, or perhaps deciding to abandon the project all together. Although, private firms are more likely to engage in FTO analysis because any infringement risk may directly affect their ability to develop new products and their ultimate profitability, public and not-for-profit private institutions are becoming increasingly aware of the need for better freedom to operate information. This is particularly true for research projects undertaken by universities or not-for-profit research centers for the specific purpose of developing new crops for developing countries. In these cases, it is critical that IP considerations be taken into account early in the research process. To date, companies and their contactors have been much more interested in FTO than non-profit institutions.

While patents are the most common type of IP right encountered, a thorough FTO analysis will assess all types of existing property rights for the likelihood that the research project or the product being commercialized infringes. Of particular concern are tangible property rights, such as cell lines, transgenic plants, germplasm, and plasmids. This is because, as described earlier, the transfer of tangible property often occurs under the terms of a material transfer agreement, which has no geographic or temporal limitation. These terms can be particularly problematic and directly impact FTO.

15.4.2. FTO Case Study: the Tomato E8 Promoter

Enabling technologies for plant transformation or transformation vectors combine several components such as promoters, selectable markers, marker removal systems, and more. Because of the fundamental role that these technologies, they have been extensively patented. In addition, the FTO surrounding plant enabling technologies is further complicated because these technologies are not used individually but are combined with a suite of related enabling technologies, specific trait technologies, and deployed in many different plant species. We can look at a relatively simple example of a single component of a transformation vector to illustrate the elements of an FTO analysis.

The target technology for this case study was a fruit-specific promoter from the tomato E8 gene. The E8 promoter has been used to improve fruit quality, extend fruit shelf life, and to express edible human vaccines specifically in ripening tomato fruit. The first step in an FTO investigation is to clearly define the target technology. In this hypothetical case, the fruit-specific promoter will be used exactly as described in the initial publications by Deikman and Fischer (1988) and Giovannoni et al. (1989). The promoters in these publications are virtually identical and consist of about 2100 nucleotides upstream of the E8 structural gene. Further promoter characterization identifying the location and sequence of functional elements within the promoter and upstream nucleotide sequence was reported in Deikman et al. (1992). These publications draw the technical boundaries surrounding the target promoter technology and provide important prior art to subsequently filed patents.

To establish the relationship of publications and patents that describe or claim the E8 promoter, a patent landscape must first be established. The patent landscape should include patents and patent applications closely related to the technology. Keywords and authors of key publications are used to search for patents or patent applications. A separate search should then be conducted to identify patents or patent applications that referenced the scientific publications describing the technology. Additionally, in the E8 case, patented DNA and protein sequence databanks were searched using the promoter's DNA sequence as a query. The patent landscape will reveal "family" relationships among different patents and published patent applications. Patent families include later patent applications that claim the benefit of an earlier, related, application or later patent applications that arise from foreign filings of the parent application. Figure 15.3 illustrates a patent family arising from a 1989 patent application related to the E8 promoter filed by Agritope: an agricultural biotechnology company.

An informative way of analyzing the FTO search results is to construct a timeline of scientific literature, patent applications, and issued patents on the specific technology and on potentially overlapping subject matter. Ordering the patents and published applications according to their priority dates (also known as effective filing dates) reveals important relationships. For example, it reveals what publications or patents are prior art against newer patents. Since patents may only be granted if the claims are both novel and nonobvious over the prior art, this analysis reveals the relative dominance of earlier, broader patents over later, narrower patents. Figure 15.4 illustrates the IP priority timeline for the E8 promoter. A thorough FTO analysis may require direct contact with the researchers and, in this analysis, it was learned from the authors of the Deikman and Fischer publication (1988) that they did not apply for patent protection prior to their publication. This information was also confirmed by searching patent databases. Based on this information, it was presumed that the basic E8 promoter technology was in the public domain, but this conclusion required thorough review and documentation of the published literature or "prior art" relative to the subject matter of subsequent patents.

As shown in the priority timeline, the Deikman and Fischer (1988) and Giovannoni et al. (1989) publications initially describe the E8 promoter technology. This precluded the novelty of any subsequent patent claims on the E8 promoter *per se* (e.g., applications filed by Agritope and Epitope). Legal counsel concluded that the tomato E8 promoter constructs *per se* can be reasonably considered to be in the public domain. However, some of the subsequent patents claim chimeric constructs comprising the E8 promoter and heterologous genes and use of the E8 promoter in conjunction with other promoter elements. Thus, certain specific uses of the E8 promoter may infringe these subsequent patents.

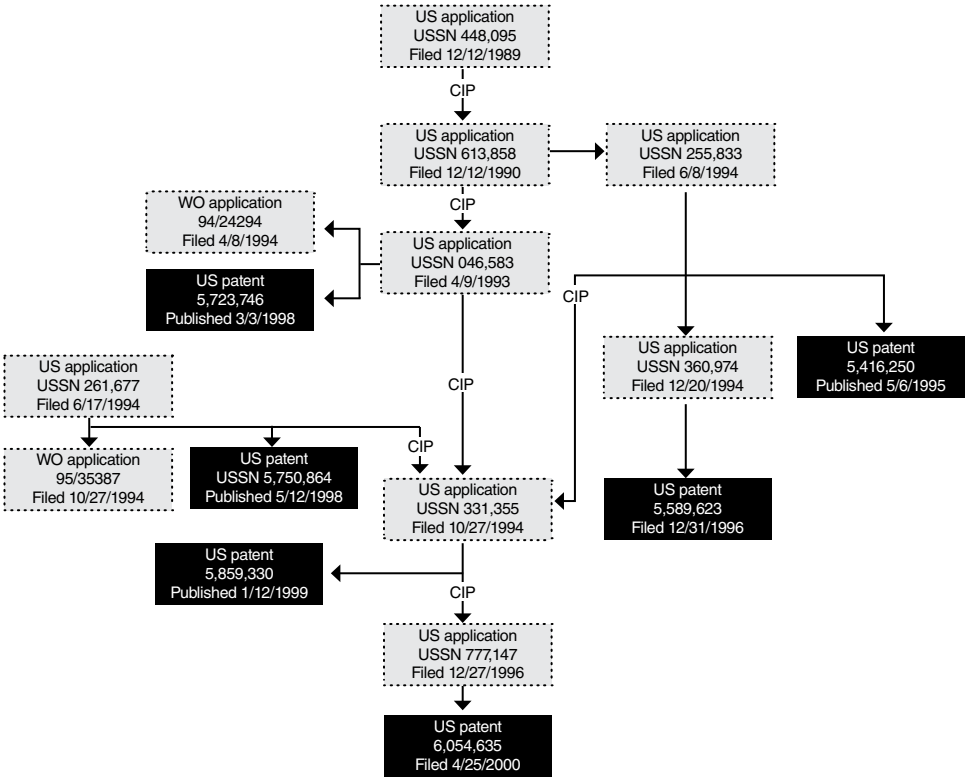


Figure 15.3. The Epitope/Agritope patent family related to the tomato E8 promoter. Patent applications are shown in gray, including US and WO or PCT applications; issued US patents are shown in black. (*Source*: Adapted from Fenton et al. (2007).)

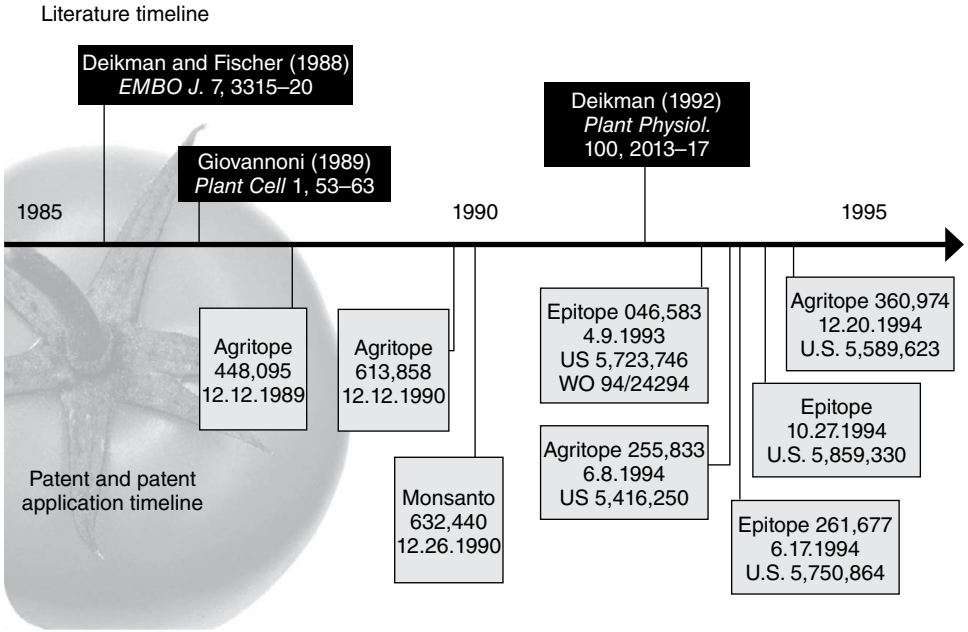


Figure 15.4. Timeline of the tomato E8 promoter scientific publications and patents. Date of publication of research articles are shown in black boxes above black timeline arrow and priority date or effective flinging date of the issued US patents associated to the tomato E8 promoter are shown in gray below the black timeline. (*Source*: Adapted from Fenton et al. (2007).)

This example provides an overview of the data and information that should be considered in an FTO analysis. It is not difficult to imagine how the complexity of an FTO analysis would grow dramatically with the inclusion of multiple enabling technologies, one or more trait technologies and proprietary germplasm. This is one of the challenges of understanding IP constraints and developing FTO strategies in plant biotechnology where multiple complementary technologies are necessarily integrated to develop new crop varieties.

15.5. STRATEGIES FOR OPEN ACCESS

The complex IP environment surrounding agricultural biotechnology research and development, exemplified by even a relatively simple FTO opinion, has spawned some new strategies and new organizations committed to lower the IP barriers to new crop development and provide more open access to patented technologies. These issues are critical for small private companies attempting to enter this sector, but can also be important for public or not-for-profit research institutions. For example, a Federal Circuit Court of Appeals ruling in the *Madey v. Duke* case emphasized that academic research is not protected by an "experimental use" exception from patent infringement, even when the research is purely fundamental (Eisenberg 2003). Most plant biotechnology laboratories routinely use patented technologies in their research without permissions. Although patent owners have rarely been concerned about academic research infringement in agriculture, there are many examples where fundamental biomedical research has been challenged because of IP issues (Marshall 2002). In addition to IP considerations in basic research, projects carried out in public or not-for-profit institutions that are targeted toward the development of crops for developing country farmers must consider the IP inputs to the project.

Most scientists are still relatively unfamiliar with how to find, understand, and utilize IP information, including published patents and patent applications. In addition, the ability to obtain the rights to use patented technologies has remained uncertain even for projects that have little commercial importance but, for example, may have large impacts in agriculture in developing countries. Several organizations have now emerged that address the relative inaccessibility of IP information and to provide a framework to ensure that IP does not block applications of agricultural biotechnology and, in particular, to facilitate projects that can have broad humanitarian benefits.

Several public sector and not-for-profit agricultural research institutions, including the University of California, the Donald Danforth Plant Science Center, North Carolina State University, Ohio State University, Boyce Thompson Institute for Plant Research, Michigan State University, Cornell University, University of Wisconsin–Madison, University of Florida, the USDA, Rutgers University, Texas A&M University, and Purdue University developed the Public Sector Intellectual Property Resource for Agriculture (PIPRA; www.pipra.org). These institutions made a public commitment to participate and promote strategies to collectively manage public-sector intellectual property in support of both US and developing country agriculture (Atkinson et al. 2003). This initial founding group of PIPRA members has grown to over 45 institutional members in 13 countries, illustrating the widespread concern and interest in working collectively to remove and avoid IP barriers that might slow development of new crops.

A number of strategies have been implemented to enhance FTO using public-sector IP for agricultural biotechnology projects. For example, informed decisions regarding dissemination of new knowledge via open publication or protecting it with a patent are clearly important and FTO can be improved if public-sector institutions systematically consider how, when, and whether to use the patent system to support broad innovation (Boettiger and Chi-Ham 2007). Even when using the patent system, PIPRA encourages its members to reserve rights to use their newest and best technologies for humanitarian purposes, particularly when they issue exclusive commercial licenses (Bennett 2007). For US agriculture, it is also important to retain rights to use patented technologies in the development of small specialty crops that are not currently within the commercial interests of

large private sector companies. The anticipated benefits of a collective IP management regime are to enable an effective assessment of FTO issues, to overcome the fragmentation of public-sector IPR and reestablish the necessary FTO in agricultural biotechnology for the public good and to enhance private sector interactions by more efficiently identifying collective commercial licensing opportunities.

Among PIPRA's core activities in developing a clearinghouse of public IP information and analytical resources, it is also developing consolidated technology packages, or patent pools, particularly in the area of enabling technologies for plant transformation. Patent pools have been used effectively by companies to expedite the development and diffusion of innovations that draw on many complementary technology components protected by multiple patents that are owned by multiple technology developers. In PIPRA's case, the development of a patent pool that provides FTO for plant transformation that incorporates patented technologies from multiple owners was facilitated by its broad membership and their shared commitment to make these technologies widely available. This strategy has created commercial licensing opportunities as well as supporting humanitarian projects.

Other initiatives, including CAMBIA, the African Agricultural Technology Foundation (AATF), and the Global Alliance for Livestock Veterinary Medicines (GALVmed), have been established to manage intellectual property issues and facilitate access to technologies in the agriculture. CAMBIA was modeled after the "open source" approach that is well developed in the IT software sector. This initiative was named "Biological Innovation for Open Society" (BiOS) and was built upon a broad philosophical foundation to "to democratize problem solving to enable diverse solutions to problems through decentralized innovation" (http://www.bios.net/daisy/bios/home.html). At the heart of BiOS was licensing language designed to preserve a pool of patented technologies from private appropriation or to create a "protected commons" of enabling agricultural biotechnologies that are made freely available and whose use cannot be restricted by third-party patent rights. Eventually, CAMBIA abandoned the BiOS model and adopted a new licensing strategy whereby it abandoned its patent and tangible property rights.

GALVmed is a nonprofit global alliance for the development of livestock vaccines to improve livelihood in developing countries. Their IP asset management strategy anticipates and mitigates IP risks throughout the value chain for the development of livestock health products (Geoola and Boettiger 2012). Similarly, AATF has implemented a public–private partnership model to access and deliver proprietary agricultural technology to small holder farmers in Africa. While many of the technologies are not patented in sub-Saharan Africa, AATF's model recognizes the value in know-how and in strategic partnerships. In 2008, AATF launched an ambitious initiative to develop water use efficient maize (WEMA). WEMA is a public/private partnership funded by the Bill and Melinda Gates Foundation, the Howard G. Buffett Foundation, and USAID. WEMA's key partners include the National Agricultural Research Institutes in Kenya, Mozambique, South Africa, Tanzania and Uganda, the International Maize and Wheat Improvement Center (CIMMYT), and Monsanto. In terms of IP, Monsanto contributed commercial drought-tolerant and insect-protection traits, royalty-free, as well as technical expertise.

15.6. CONCLUSIONS

Intellectual property as a tool to foster innovation has been important for over two centuries but has become a much more prominent feature of research in the life sciences and in agricultural biotechnology, in particular, in just the last 25 years. This trend is unlikely to be reversed and, indeed, the importance of intellectual property as an intangible asset contributing to the value of life science companies continues to increase. However, robust and sustained innovation in agricultural biotechnology, as in many technology sectors, requires a balance of both exclusive and nonexclusive access to proprietary technologies. This balance should ensure that the fundamental research tools are broadly available to support research and development in many application areas and at the same time provide the exclusivity to specific trait or trait/crop combinations that will allow the developer

of new varieties to recoup their substantial investment. The public sector has a role to play alongside agricultural biotechnology companies, particularly in providing research tools and broad enabling technologies and in addressing biotechnology applications in specialty crops whose market size may not justify commercial investment. The most powerful approaches, however, will come from public–private partnerships that mobilize proprietary technologies to address agricultural biotechnology product developments that have a high social but low commercial value, including strategies to feed some of the world's poorest populations.

LIFE BOX 15.1. ALAN BENNETT

Alan Bennett, Executive Director of Public Intellectual Property Resource for Agriculture (PIPRA), Professor of Plant Sciences, University of California, Davis.



Alan Bennett. Courtesy of Alan Bennett.

My professional formative years were as an undergraduate in college during the 1970s. There were a couple of important events that shaped my professional view and aspirations. The first was when I took some time off to work as a laborer on the Alaska pipeline that was under construction at the time. It was a dream job because it paid well with lots of overtime hours. For me, it had two impacts. The first was that it allowed me to afford to complete my college education, but more importantly, it firmly convinced that I wanted my life's work to be both interesting and in some small way to help make the world a better place. At the same time, there was a growing concern about food security in the world and I became convinced that fundamental research in plant biology could play a role in feeding the growing global population, and I remain convinced of this today.

After graduate school at Cornell University, I was lucky enough to get an assistant professor appointment at my alma mater, the University of California-Davis, where I have spent my entire professional life. I started my career in the early days of plant molecular genetics and genetic engineering and recognized that these new tools could be the vehicle to link fundamental research in plant biology to applications for global food security. It was also becoming clear that intellectual property and patents were important tools in making this linkage between fundamental science and applications, and both companies and universities were getting more active in protecting research results—even very fundamental research results. This became an intense interest of mine and I was able to move sideways from my professor position to the job of managing the technology transfer program for the University of California system. I spent nearly a decade working in this area. However, despite good intentions, universities were not experienced in using intellectual property (IP) strategically and as a consequence there were a few high profile cases where university patents were hindering, rather than promoting, global food security research applications. This situation caught the attention of Gary Toenniessen at the Rockefeller Foundation who began a dialog among US universities encouraging them to think about strategies to manage agricultural IP with the intention of supporting both commercial applications as well as humanitarian applications in support

of food security. This dialog resulted in the formation of the Public Intellectual Property Resource for Agriculture (PIPRA), and UC-Davis successfully competed to host this organization and I competed to be its director. PIPRA has worked as both a watchdog and a resource to promote the strategic use of IP in agriculture by supporting research programs with analysis and advice and by providing developing-country technology managers with the skills and tools they need to work effectively in this domain. In just a few years, PIPRA has provided educational programs to over 200 technology managers from over 40 developing countries.

From a personal perspective, the ability to move between roles as a teacher, researcher, and an intellectual property policy advocate has certainly met my early goal of having my life's work be both interesting and in some small way making the world a better place. I can attribute this to two indispensible elements. The first is the good fortune to work for a public land-grant university that itself is fundamentally committed to improving society and the world and that is large enough to have provided me a number of diverse professional opportunities throughout my career. The second, of course, is personal motivation and the ability to build partnerships among like-minded people and institutions. PIPRA is a great example of forging partnerships with, first, the Rockefeller Foundation and then, later, with many like-minded universities to work toward a common goal. For me, training in science and research was a great start in my professional life because it gave me analytical skills and deep knowledge in one topic—but it also gave me the opportunities to contribute in other ways to the policies and strategies that are critical to the applications of science. It has all been interesting. I'm really not sure what opportunities lie around the corner but, if I can judge the future by the past, I am definitely looking forward to what is next.

LIFE BOX 15.2. MAUD HINCHEE

Maud Hinchee, Chief Science Officer of Agricen Sciences.



Maud Hinchee. Courtesy of Maud Hinchee.

I certainly did not plan on being a plant biologist. However, my mother always thought I would be a botanist, because I eliminated her eggplant yield from her backyard garden by sterilizing the flowers without her knowing it (I hated eggplant). It wasn't until I took a college course that captured my imagination that I decided to become a botanist. The class was plant anatomy, which in some universities can be quite dry. However, this course was taught by Dr. Tom Rost at the University of California, Davis—a young professor who taught using an experimental approach to understanding the form, structure, and function of plant cells, tissues, and organs, while allowing us to appreciate the esthetic beauty of plant cells. As a somewhat artistic type, I liked this blend of scientific discovery and microscopic art. I went on to receive my BS degree in botany from the University of California, Davis, (UCD) in 1975, and then my MS in botany from

the University of Washington (UW). In my undergraduate research, I studied the development of roots. At UW, I compared and contrasted the anatomical and growth characteristics between aerial and soil roots of *Monstera deliciosa*. I returned to UCD where I received my PhD degree in 1981 in plant morphogenesis. My project was to determine what effect the cotyledons of pea had on the development and distribution of lateral roots in young seedlings.

How these various research projects enabled me to become a plant biotechnologist is probably a matter of being in the right place at the right time. Researchers were just starting to make some headway in developing methods for inserting genes into plant cells. Since so little was known at that time as to what controlled which cells successfully incorporated DNA, and which of these cells subsequently could develop into a whole plant, I was able to provide valuable insights to the process as a plant morpho-geneticist. I did my first training in plant transformation techniques during a postdoctoral research associate position at the Hawaiian Sugar Planters' Experiment Station in Hawaii, working on the incorporation of DNA into sugarcane protoplasts.

I then was hired at Monsanto Co, in St. Louis, Missouri, in 1982. My first role was to determine why regeneration and transformation experiments in soybean weren't leading to the expected results. This activity provided me much insight into the cellular basis for the regeneration process and allowed me to design methods to specifically target our genetic engineering tool, *Agrobacterium*, to the right cells at the right time. The result was a successful and reproducible soybean transformation protocol that yielded the first transgenic soybean containing the Roundup Ready gene. Today, 90% of the soybeans grown in the United States have this trait. It gives me great pride still to drive by a soybean field that is clean of weeds and realize that the resulting benefit to farmers is due, in some small way, to my research efforts. Working at Monsanto was the greatest learning experience of my life. Besides the opportunity to develop transformation methods for a variety of crops that included sugar beet, flax, potato, strawberry, cotton, and sweet potato, I also learned how a biotech product was "built" from the ground up—from conception of the gene construct all the way through to regulatory approval of a transgenic plant. Another rewarding experience I had in this time was leading a team of Monsanto and African scientists to develop virus-resistant sweet potato for subsistence farmers in Kenya and other parts of Africa. All this experience served me well, in my next role as a the technical lead for a business team dedicated to developing biotech collaborations in specialty crops worldwide in crops such as forestry, sugarcane, and fruits and vegetables.

I left Monsanto in 2000 to become the Chief Technology Officer of ArborGen LLC, a forestry biotechnology company that currently develops genetically improved planting stock for the pulp, timber, and bioenergy industries. As much as I enjoyed my time at Monsanto, I enjoyed guiding a young company toward successful product development. I foresee in the future that transgenic technologies will play a very important role in sustaining our environment by providing solutions to the worsening energy crisis. ArborGen will be marketing trees that will require a relatively small land "footprint" because of the trees' high productivity and which can supply a renewable and sustainable source of biomass for the production of cellulosic ethanol.

In 2013, I became the Chief Science Officer of Agricen Sciences, which is at the forefront of developing new types of biological technologies that function as biostimulants to enhance plant growth, yield, and health. These technologies fit into a systems-based approach to manage crop production sustainably. The systems-based approach capitalizes on newly emerging knowledge of how the genetics and physiology of a crop interacts with its own microbiome, and that of the soil, in the context of modern agricultural systems. I am very proud that my career has continued to evolve in conjunction with new concepts and technologies.

REFERENCES

An YQ, McDowell JM, Huang S, McKinney EC, Chambliss S, Meagher RB (1996): Strong, constitutive expression of the Arabidopsis ACT2/ACT8 actin subclass in vegetative tissues. *Plant J* **10**:107–121.

Atkinson RC, Beachy RN, Conway G, Cordova FA, Fox MA, Holbrook KA, Klessig DF, McCormick RL, McPherson PM, Rawlings HR, Rapson R, Vanderhoef LN, Wiley JD, Young CE (2003): Public sector collaboration for agricultural IP management. *Science* **301**:174–175.

Barton KA, Binns AN, Chilton M-D, Matzke AJM (2000): Regeneration of plants containing genetically engineered T-DNA. U.S. Patent 6,051,757.

Beachy R, Bhattacharyya M (1998): Plant promoter. U.S. Patent 5,824,857.

Bednarek SY, Wilkins TA, Dombrowski JE, Raikhel NV (1989): A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. *Plant Cell* **2**:1145–1155.

Benfey PN, Chua N-H (1992): Plant promoters. U.S. Patent 5,110,732.

Bennett AB (2007): Reservation of rights for humanitarian use. In *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*, Krattiger A, Mahoney R, Nelsen L, Thomson J, Bennett A, Satyanarayana K, Graff G, Fernandez C, Kowalski S (eds). MIHR/PIPRA, Oxford/Davis, CA.

Blechl A, Anderson O, Somers DA, Torbert KA, Rines HW (1999): Glutenin genes and their uses. U.S. Patent 5,914,450.

Boettiger S, Chi-Ham C (2007): Defensive publishing and the public domain. In Krattiger A, Mahoney R, Nelsen L, Thomson J, Bennett A, Satyanarayana K, Graff G, Fernandez C, Kowalski S (eds). *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*. MIHR/PIPRA, Oxford/Davis, CA.

Bojsen K, Donaldson I, Haldrup A, Joersboe M, Kreiberg JD, Nielsen J, Okkels FT, Petersen SG (1998): Mannose or xylose based positive selection. U.S. Patent 5,767,378.

Clark D, Klee H, Dandekar A (2004): Despite benefits, commercialization of transgenic horticultural crops lags. *Cal Agric* **58**:89–93.

Cohen SN, Boyer, HW (1980): Process for producing biologically functional molecular chimeras. U.S. Patent 4,237,224.

Comai L, Sanger MP, Daubert SD (2000): Figwort mosaic promoter and uses. U.S. Patent 6,051,753.

Conkling MA, Mendu N, Song W (1998): Root cortex specific gene promoter. U.S. Patent 5,837,876.

Conway G, Toenniessen G (1999): Feeding the world in the twenty-first century. *Nature* **402**:C55–58.

Dehesh K, (2002): Plastid transit peptide sequences for efficient plastid targeting. U.S. Patent 2002,0178,467.

Deikman J, Fischer RL (1988): Interaction of a DNA binding factor with the 5'-flanking region of an ethylene-responsive fruit ripening gene from tomato. *EMBO J* **7**:3315–3320.

Diekman J, Kline R, Fischer RL (1992): Organization of ripening and ethylene regulatory regions in a fruit-specific promoter from tomato (*Lycopersicon esculentum*). *Plant Physiol* **100**:2013–2017.

Dirk LM, Williams MA, Houtz RL (2001): Eukaryotic peptide deformylases. Nuclear-encoded and chloroplast-targeted enzymes in Arabidopsis. *Plant Physiol* **127**:97–107.

Dirk LM, Williams MA, Houtz RL (2002): Specificity of chloroplast-localized peptide deformylases as determined with peptide analogs of chloroplast-translated proteins. *Arch Biochem Biophys* **406**:135–141.

Dobres, MS Mandaci. 1998. Plant promoter useful for directing the expression of foreign proteins to the plant epidermis. US Patent 5,744,334

Eisenberg R (2003): Patent swords and shields. *Science* **299**:1018–1019.

Feldman M (2005): Commercializing Cohen-Boyer 1980-1997. Rotman School of Management, University of Toronto, http://sites.kauffman.org/pdf/tt/Feldman_Maryann.pdf

Fenton G, Chi-Ham C, Boettiger, S (2007): Freedom to operate: the law firms approach and role. In Krattiger A, Mahoney R, Nelsen L, Thomson J, Bennett A, Satyanarayana K, Graff G, Fernandez C, Kowalski S (eds). *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*. MIHR/PIPRA, Oxford/Davis, CA

van der Fits L, Memelink J (1997): Comparison of the activities of CaMV 35S and FMV 34S promoter derivatives in *Catharanthus roseus* cells transiently and stably transformed by particle bombardment. *Plant Mol Biol* **33**:943–946.

Fraley RT, Horsch RB, Rogers SG (1991): Genetically transformed plants. EP0131620.

Gelvin SB, Hauptmann R, Ni M, Cui D (1999): Chimeric regulatory regions and gene cassettes for expression of genes in plants. U.S. Patent 5,955,646.

Geoola J, Boettiger S (2012): *Intellectual Property Management in Livestock Veterinary Medicines for Developing Countries. Agricultural Innovation Systems: An Investment Sourcebook*. The World Bank: International Bank for Reconstruction and Development/International Development Association, Washington, DC.

Giovannoni J, DellaPenna D, Bennett AB, Fischer RL (1989): Expression of a chimeric polygalacturonase gene in transgenic rin (ripening inhibitor) tomato fruit results in polyuronide degradation but not fruit softening. *Plant Cell* **1**:53–63.

Graff GD, Cullen SE, Bradford KJ, Zilberman D, Bennett AB (2003): The public–private structure of intellectual property ownership in agricultural biotechnology. *Nat Biotechnol* **21**:989–995.

Graff, G. D., Zilberman, D. Bennett, AB (2009): The contraction of agbiotech product quality innovation. *Nat Biotechnol* **27**:702–704.

Graff GD, Phillips D, Lei Z, Oh S, Nottenbur C, Pardey PG (2013): Not quite a myriad of gene patents. *Nat Biotechnol* **31**:404–410.

Grundler F, Nitz I, Puzio P (2001): Root-specific promoter. PCT WO 01/44454.

Hamilton CM (1998): Binary BAC vector. U.S. Patent 5,733,744.

Harada JJ, Lotan T, Ohto M-A, Goldberg RB, Fischer RL (2001): Leafy cotyledon genes and their uses. U.S. Patent 6,320,102.

Hardin G (1968): Tragedy of the commons. *Science* **162**:1243–1248.

Hayashi M, Aoki M, Kato A, Kondo M, Nishimura M (1996) Transport of chimeric proteins that contain a carboxy-terminal targeting signal into plant microbodies. *Plant J* **10**:225–234.

Heller MA, Eisenberg RS (1998): Can patents deter innovation? The anticommons in biomedical research. *Science* **280**:698–701.

Herrera-Estrella L, Depicker A, Van Montagu M, Schell J (1983): Expression of chimeric genes transferred into plant cells using a Ti plasmid-derived vector. *Nature* **303**:209–213.

Herrera-Estrella L, VanDen Broeck G, Van Montagu M, Scheier P, Schell J, Bohnert HJ, Cashmore AR, Timko MP, Kausch AP (2000): Chimaeric gene coding for a transit peptide and a heterologous peptide. U.S. Patent 6,130,366.

Hoekema A, Hooykaas PJJ, Schilperoort RA (1992): Process for introducing foreign DNA into the genome of plants. U.S. Patent 5,149,645.

Huang S, An Y Q, McDowell JM, McKinney EC, Meagher RB (1997): The Arabidopsis ACT11 actin gene is strongly expressed in tissues of the emerging inflorescence, pollen, and developing ovules. *Plant Mol Biol* **33**:125–139.

Jefferson T (1987): Thomas Jefferson to Isaac McPherson In: *The Founders' Constitution*, Kurland PB, Lerner R (eds). University of Chicago Press. Available at: http://press-pubs.uchicago.edu/founders/documents/a1_8_8s12.html (accessed on September 13, 2015).

Kato A, Hayashi M, Kondo M, Nishimura M (1996): Targeting and processing of a chimeric protein with the N-terminal presequence of the precursor to the glyoxysomal citrate synthase. *Plant Cell* **8**:1601–1611.

Klein TM, Wolf ED, Wu R, and Sanford JC (1987): High velocity microprojectiles for delivering nucleic acids into living cells. *Nature* **327**:70–73

Komarnytsky S, Borisjuk NV, Borisjuk LG, Alam MZ, Raskin I (2000): Production of recombinant proteins in tobacco guttation fluid. *Plant Physiol* **124**:927–934.

Kowalski SP, Ebora RV, Kryder D, Potter RH (2002): Transgenic crops, biotechnology and ownership rights: what scientists need to know. *Plant J* **31**:407–421.

Kryder RD, Kowalski SP, Krattiger AF (2000): *The intellectual and technical property components of pro-vitamin A rice (GoldenRice™): A preliminary freedom-to-operate review*. Brief No. 20. ISSAA, Ithaca, NY.

Maiti I, Shephard RJ (1998): Promoter (FLt) for the full length transcript of peanut chlorotic strek caulimovirus (PCLSV) and expression of chimeric genes in plants. U.S. Patent 5,850,019.

Maiti I, Shephard RJ (1999): Full length transcript (FLt) promoter from figwort mosaic caulimovirus (FMV) and use to express chimeric genes in plant cells. U.S. Patent 5,994,521.

Marshall E (2002): DuPont ups the ante on use of Harvard's OncoMouse. *Science* **296**:1212–1213.

McElroy D, Wu R (2002): Rice actin 2 promoter and intron and methods for use thereof. U.S. Patent 6,429,357.

McElroy D, Zhang W, Cao J, Wu R (1990): Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell* **2**:163–171.

Mentewab A, Stewart CN Jr (2005): Overexpression of an *Arabidopsis thaliana* ABC transporter confers kanamycin resistance to transgenic plants. *Nat Biotechnol* **23**:1177–1180.

Miki B, McHugh S (2004): Selectable marker genes in transgenic plants: applications, alternatives and biosafety. *J Biotechnol* **107**:193–232.

Odell JT, Nagy F, Chua N-H (1985): Identification of DNA sequences required for activity of a plant promoter: the CaMV 35S promoter. *Nature* **313**:810–812.

Pray CE, Naseem A (2005): Intellectual property rights on research tools: incentives or barriers to innovation? Case studies of rice genomics and plant transformation technologies. *AgBioForum* **8**(2&3), Article 7.

Roa-Rodriguez C, Nottenburg C (2003): *Antibiotic Resistance Genes and Their Uses in Genetic Transformation*. CAMBIA IP Resource, Canberra. Available at: www.cambiaip.org (accessed on September 13, 2015).

Rogers SG, Fraley RT (1993): Chimeric genes suitable for expression in plant cells, method for selection of transformed plant cells. U.S. Patent 08/127100.

Rogers SG, Fraley RT (2001): Chimeric genes suitable for expression in plant cells. U.S. Patent 6,174,724.

Rogers SG, Horsch RB, Fraley RT (2012): Genetically transformed plants. U.S. Patent 8,273,954.

Romano CP, Cooper ML, Klee HJ (1993): Uncoupling auxin and ethylene effects in transgenic tobacco and Arabidopsis plants. *Plant Cell* **5**:181–189.

Rudder S, Doohan F, Creevey CJ Wendt T, Mullins E, (2014): Genome sequence of *Ensifer adhaerens* OV14 provides insights into its ability as a novel vector for the genetic transformation of plant genomes. *BMC Genomics* **15**:268.

Sanford JC, Wolf ED, Allen NK (1991): Method for transporting substances into living cells and tissues and apparatus therefore. U.S. Patent 5,036,006.

Santerre RF, Rao RN (1988): Recombinant DNA cloning vectors and the eukaryotic and prokaryotic transformants thereof. U.S. Patent 4,727,028.

Smeekens S, Bauerle C, Hageman J, Keegstra K, Weisbeck P (1986): The role of the transit peptide in the routing of precursors toward different chloroplast compartments. *Cell* **46**:365–375.

Tague BW, Dickinson CD, Chrispeels MJ (1989): A short domain of the plant vacuolar protein phytohemagglutin targets invertase to the yeast vacuole. *Plant Cell* **2**:533–546.

Volokita M (1991): The carboxy-terminal end of glycolate oxidase directs a foreign protein into tobacco leaf peroxisomes. *Plant J* **1**:361–366.

Wendt T, Doohan F, Mullins F (2012): Production of Phytophthora infestans-resistant potato (*Solanum tuberosum*) utilising Ensifer adhaerens OV14. *Transgenic Res* **21**:567–578.

Ye XD, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000): Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**:303–305.

Yoder JI, Lassner M (1998): Biologically safe plant transformation system. U.S. Patent 5,792,924.

CHAPTER 16

Why Transgenic Plants Are So Controversial

JENNIFER TRUMBO

Department of Nutrition, University of Tennessee, Knoxville, Tennessee

DOUGLAS POWELL

Brisbane, Australia

16.0. CHAPTER SUMMARY AND OBJECTIVES

16.0.1. Summary

In many fields, controversy remains over the role of science in the environment and peoples' lives. Plant biotechnology has seemingly incited much protest in its relatively short commercial lifetime. Other than the scientific reasons (risk assessment) given in Chapter 14, there are many aspects of plant biotechnology that encourage debate. Societal, political, and economic factors contribute to this dynamic. In addition, Mary Shelley's *Frankenstein* illustrates the roles of fear, philosophy, and religion in the controversy. While innovation invokes change, and change should be assessed for risks and benefits, fear has seemingly oftentimes dominated arguments and undermined the risk assessment process. Although the debate over genetically modified organisms (GMOs) continues, scientists and citizens can take steps to understand and balance risk with benefits through education and communication.

16.0.2. Discussion Questions

1. Why is Frankenstein's monster ("Frankenfoods") often used to illustrate the risks of biotechnology?
2. What are some of the factors that play into peoples' perception of risk?
3. What are the stigmas associated with plant biotechnology, and how can they be overcome?
4. What two major scientific stories prompted media attention toward transgenic plants in the 1990s?
5. What issues are still being debated? Should they be?
6. How does the private sector influence public perception of plant biotechnology risks?
7. What can the scientific community do in the future to encourage communication among all sides of the controversy?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

16.1. INTRODUCTION

Carl Sagan once said "We live in a society exquisitely dependent on science and technology, in which hardly anyone knows anything about science and technology (1989)." Science can be very intimidating for those outside of its reach. It can be quite a challenge even for seasoned scientists to keep up with the latest developments in their fields. We live alongside the results of countless years of scientific development. Many of these inventions we accept even though we do not fully understand them. I admit that I do not know every mechanism that allows my car to deliver me to work each and every day. I acknowledge that one day I may face some risk because of this technology, such as brake failure, yet I continue to use it because walking to work would take hours. The benefits outweigh the risks. When scientific innovation intersects with daily life, it can be difficult to know how to react. Combine this with environmental, health, economic, and political concerns and controversy is born.

Today, many farmers around the world embrace scientific advancements such as agricultural biotechnology. It seems that comparatively fewer consumers understand the science and regulation of products resulting from agricultural biotechnology. This knowledge gap has spurred the development of environmental and activist groups that lobby for greater regulation and labeling of GM products. They, along with Michael Crichton ("Jurassic Park" 1990), find biotechnological innovations to be scary and in need of special scrutiny. This chapter discusses factors in the development of controversies over GMOs derived through agricultural biotechnology. While the controversies and tensions between society and science have lasted for centuries, industry and especially regulators and public scientists can address many of society's issues through transparency, communication, and education.

16.1.1. The Frankenstein Backdrop

First published in 1817, Mary Shelley's *Frankenstein* described the dangers of unregulated science. Many scientists at that time were trying to discover how to create life. Shelley explored the ramifications of this search in the following passage:

> The ancient teachers of this science promised impossibilities, and performed nothing. The modern masters promise very little; they know metals cannot be transmuted, and that the elixir of life is a chimera. But these philosophers, whose hands seem only made to dabble in dirt, and their eyes to pore over the microscope or crucible, have indeed performed miracles. They penetrate into the recesses of nature, and show she works in her hiding places. They ascend into the heavens; they have discovered how the blood circulates, and the nature of the air we breathe. They have acquired new almost unlimited powers; they can command the thunders of the heaven, mimic the earthquake, and even mock the invisible world with its own shadows (p. 37).

The above passage describes the duality of scientific knowledge. Our understanding of scientific discoveries is limited by human nature. We can predict some potential benefits of scientific advancements, but the full consequences, especially the negative aspects of innovation, might be downplayed. Frankenstein shows the role of arrogance in the misuse of scientific knowledge. This story resonates with the common fear of science out of control. Much of the controversy over genetic modification is born from this fear. Activist groups exploit this by labeling GM foods "frankenfoods." Science and technology has progressed despite naysayers. In agriculture, many innovations have laid the foundation for genetic modification of plants; its dissenters follow.

16.1.2. Agricultural Innovations and Questions

Chemicals have been used in agricultural food production for centuries. As early as 1000 BC, Chinese farmers used sulfur as a fumigant. In the sixteenth century, arsenic-containing compounds were utilized as insecticides, and by the 1930s the production of modern synthetic chemicals commenced.

The onset of World War II encouraged the rapid production and use of chemical substances such as DDT, which was used for control of malaria-transmitting insects. The postwar era marked the start of the modern agrochemical industry. As a direct result of technical advancements in chemical production during this period, various insecticides, fungicides, and fumigants found their place in agriculture and food production (Powell and Leiss 1997).

Today, as an alternative to chemical application, genes from naturally occurring organisms with pesticidal properties are being engineered into plants. This is advantageous because broad spectrum pesticides, while addressing practical pest issues, carry risks for pesticide applicators and the environment. Yet, the public discussion of agricultural biotechnology has been framed narrowly in terms of risks versus benefits of single genes while generally not fully considering entire agricultural systems. This perspective ignores the ultimate objective: to maximize benefits in complete perspective of societal and environmental welfare.

The public discussion of genetically engineered foods has, since at least 1998, been characterized by seemingly simple questions that many advocates have failed to adequately answer, including the following: Why are you messing with nature? Why don't you label everything? Can you guarantee there won't be any long-term risks? Why are you playing God?

A May 2007 review of a documentary film, *The Future of Food* (available at http://www.newstarget.com/021827.html), while exaggerated, summarizes much of the concern regarding genetically engineered food:

> There is a cabal of power-hungry corporations that are systematically destroying humanity's future. These companies have taken over the food supply, injected pesticides, viruses and invading genes into staple crops, engineered "terminator" genes that make crop seeds unviable, destroyed the livelihood of farmers and used every tactic they could think of—legal threats, intimidation, bribery, monopolistic market practices and many more—to gain monopolistic control over the global food supply. One documentary brings you this astonishing story. Through the testimony of family farmers, ecological scientists, agricultural experts and numerous public documents, *The Future of Food* tells a horrifying, heart-stopping story of how Big Agriculture has sold out the future of human civilization for the almighty dollar.

"Big Agriculture" has been selling the products of agricultural biotechnology since the US introduction of the Flavr Savr tomato in 1994. Molecular biology tools move and alter specific genes to bolster crop productivity, extend the shelf life of fresh fruits and vegetables, and reduce the effect of environmental stresses inherent in food production.

Since 1995, farmers from North America and across the globe decided that paying technology fees and other costs associated with premium corn, soybean, canola, cotton, and other crop seeds was worthwhile because the new traits increased their profit. Genetically enhanced crops generally had increased yields using the same amount of land, reduced chemical, and labor costs and increased farming system efficiency. So, why has this technology engendered such deep hostility? As discussed in *The Future of Food* (2007), some people think GMOs are risky, ill-founded, and allow large companies to have excessive control over the global food system.

16.2. PERCEPTIONS OF RISK

Exactly how an individual perceives and measures risk—in this case the risk posed by genetically engineered food—has been the subject of extensive research. Sandman (1987) noted that the public generally pays too little attention to the hazardous nature of risks, and experts usually completely ignore those factors that fuel consumer unrest or outrage. Scientists, in general, define risks in the language and procedures of science itself; they consider the nature of the harm that may occur, the probability that it will occur, and the number of people who may be affected (Groth 1991). Most

consumers seem to be less aware of the quantitative or probabilistic nature of a risk, and appear to be much more concerned with broader, qualitative attributes, such as the following:

- If the risk is voluntarily assumed
- If the risks and benefits are fairly distributed
- If the risk can be controlled by the individual
- If a risk is necessary and unavoidable
- If there are safer alternatives
- If the risk is familiar or exotic
- If the risk is natural or technological in origin

The list above can be expanded (Sandman 1987). While such generalizations are of limited value, they can greatly influence public opinion and behavior. For example, people generally accept the risks of driving as necessary and unavoidable, although there were 10.8 million traffic accidents in the United States from 1990 to 2009 according to the US Census Bureau. Compare this value to the substantiated injury count of GMOs because of the agricultural biotechnology techniques used: 0 (Key et al. 2008).

According to Covello (1983, 1992), psychological sciences research has identified 47 known factors that influence risk perception including control, benefit, and, the most important factor, trust. While these factors can help explain why consumers are concerned about a potential risk such as consuming food made from genetically engineered plants, differences in risk perception only superficially explain the visceral outrage that has greeted genetically engineered crops in some areas. By examining the various social actors and their tactics of public persuasion, a general picture emerges that helps explain the social controversy surrounding genetically engineered crops.

One factor in the controversy over GM foods rests upon authority figures. Those individuals responsible for food safety risk management ought to be viewed in the light of reducing, mitigating, and/or minimizing a particular risk. Those responsible must be able to effectively communicate their efforts and to demonstrate they are actually reducing levels of risk. Sometimes, communication is not effective and transparent enough for the public to understand and trust. In turn, we have seen stigma grow in relation to GM foods worldwide. Later in this chapter, we will specifically discuss the corporate sector's influence on public opinion. First, we will explore, in general, how public opinion is formed.

Stigma is a powerful shortcut that consumers may use to evaluate foodborne risks. Gregory et al. (1995) have characterized stigma as follows:

- The source is a hazard
- A standard of what is right and natural is violated or overturned
- Impacts are perceived to be inequitably distributed across groups
- Possible outcomes are unbounded (scientific uncertainty)
- Management of the hazard is questionable

These factors certainly apply to agricultural biotechnology products. Stigmatization is becoming the norm for food and water linked to human illness or even death. Obesity, one of the United States's most prominent health problems, is linked to food quality. Yet, many individuals continue to focus on issues like agricultural biotechnology rather than the level of processing or amount or form of sugars in what we eat every day because this is not stigmatized to the same degree. The challenge then is, how to reduce stigma? Potential solutions include the following factors:

- Effective and rapid surveillance systems
- Effective communication about the nature of risk

- A credible, open, and responsive regulatory system
- Demonstrable efforts to reduce levels of uncertainty and risk
- Evidence that actions match words

Appropriate levels of risk management coupled with sound science and excellent communication about the nature of risk are required to further garner the benefits of any technology, including agricultural biotechnology.

The products of agricultural biotechnology became mainstream around the same time that the North American public was being exposed to massive amounts of microbial food safety information, beginning with the Jack-in-the-Box restaurant *Escherichia coli* O157:H7 outbreak of 1993 (Powell and Leiss 1997). This led to unprecedented interest in the way food is produced. Consumer concerns about food safety issues—such as mad cow disease, *E. coli* O157:H7, and salmonella—have been pushed from the supermarket all the way back to the farm, such that any and all agricultural practices are now coming under public scrutiny. This trend is reflected by increased organic food sales, books like the *100-Mile Diet* (Smith and Mackinnon 2007) and the growth of community-shared agriculture (CSA), as individuals seek to exert more control over the food that nourishes their bodies and souls. Trends toward vegan, vegetarian, free-range, and free-trade labels demonstrate the growing consumer interest in "natural" foods, although processed foods are still very popular and more cost effective for people of lower incomes. In many ways, products from agricultural biotechnology have been lumped in with processed foods as unhealthy and unnatural. Upon visiting a grocery store, you will likely see "vegan," "fat-free," "whole grains," and other health-associated labels alongside "non-GMO" labels on food products. While these products might be more healthful than their alternatives, no one has demonstrated direct positive health properties from the lack of GMO ingredients.

Admittedly, fears regarding food safety were not unfounded. During the mid-1990s, several industry and government practices promoted suspicion. The results of GE crops field trials were often difficult for the public to obtain, which created an atmosphere of distrust. Industry groups often argued that genetic engineering of crops was an extension of bread- and winemaking in an attempt to make the unfamiliar familiar. John Durant has noted that attempts to characterize biotechnology as merely trivial extensions of the familiar techniques of baking, viniculture, and breeding are "pedantic" at best: "The technologies employed are completely different and it is the power and precision of the new molecular biology that drives both industrial growth and public concern" (Durant 1992). Comparisons to traditional breeding tend to magnify rather than soothe consumer concerns no matter their validity.

Individuals supportive of genetically engineered crops argued that the term "genetic engineering" was alarming to the public, and instead terms like "crop improvement" and "biotechnology" should have been used. Activists responded with terms like "Frankenfood" and "GMO," which now dominate public discourse. The topic of genetically engineered food was endlessly surveyed around the world, with public notions of agricultural biotechnology consistently articulated as concerns about uncertainty, "playing God," and the involvement of powerful industry interests.

16.3. RESPONSES OF FEAR

In response to controversies like agricultural biotechnology, many politicians, company executives, and academics urge citizens to become better educated in all matters scientific, to therefore overcome public fear as a barrier to "progress." This rhetorical strategy has been advocated by technology promoters in discussions of technological risk for the past 200 years. More recently, promoters of agricultural chemicals in the 1960s and nuclear energy in the 1970s embraced the public education model. Today, the notion of public education is the basis of dozens of communication strategies supported by government, industry, and scientific societies,

in the absence of any data suggesting that such educational efforts are successful. As noted by Kelley (1995), voters in democracies routinely make decisions about policies about which they have no detailed academic understanding. Consumers have and will continue to make decisions about genetically engineered foods, whether they are "better educated" or not. Fear is a compelling motivator and strong adversary, which education has a difficult time defeating. While education can be effective, a party must be willing to realize and engage the information at hand.

Genetic engineering is a powerful technology. Gupta et al. (2012) suggest that all revolutionary technologies create three public responses, in succession: unrealistic expectations (all new technologies are oversold; there is an old saying that "bullshit is the grease on the skids of innovation"), confusion, and generally finding a way to cope. Biotechnology has been, and continues to be, oversold; but as with other new technologies, a public discussion over time shifts from one of risks versus benefits to a more realistic approach of extracting whatever benefits a technology can bring while actively and prudently minimizing risks. But in many areas of the world, particularly Europe, public discussion of genetically engineered foods remains in the early stages. European Commission reports from 1985 to 2010 concluded that GM plants did not pose a relevant health risk compared to traditional products. Yet, this authority continues to not clearly communicate their findings. Consequently, the products are still stigmatized. This pattern is not representative of other controversial issues. For example, in France, nuclear energy is embraced. In the United States, mentioning nuclear energy is still very much a political faux pas. The perception of risk involved in nuclear power is therefore much different in France than in the United States.

This suggests that the varying degrees of public controversy in social groups across the globe is directed by repetitious conversations (including media coverage) about risks and benefits, what one is exposed too, and what one chooses to acknowledge. It is therefore a sociological phenomenon where similar information can produce completely different results. Many factors have produced the GMO controversy particular to the United States. The advent of social networking has increased information flow and communal communication among a wide sector of society. Wikipedia, YouTube, Facebook, and other Internet-based social networking websites produce a unique type of information democratization. Therefore, even the most scientifically flawed and unsubstantiated information can spread among the population in a moment. It takes much longer than a moment to remediate this misinformation. Groups like the Genetic Literacy Project (http://www.geneticliteracyproject.org/) promote scientific literacy. Yet, there are several misguided stories have shaped the current view of genetically engineered foods. A few relevant stories from the United States and Canada are dissected in the text that follows.

In 1989, there was an outbreak in the United States of a newly recognized fatal blood disease called *eosinophilia–myalgia syndrome* (EMS). The outbreak killed at least 27 people and sickened another 1500. After extensive investigation, the cause was finally traced to certain batches of the amino acid L-tryptophan, manufactured in Japan by Showa Denko and widely available in the United States as a nutritional supplement. It has been estimated that, prior to this outbreak, ≤2% of the US population took L-tryptophan to manage insomnia, premenstrual syndrome, stress, and depression in the absence of any medical data supporting the effectiveness of the supplement. In 1991, L-tryptophan was banned in the United States. While these restrictions have relaxed in the past decade, importation is still blocked.

L-Tryptophan is manufactured in a fermentation process using a bacterium, *Bacillus amyloliquefaciens*, in the same way that yeast ferments the sugars in barley into ethanol in beer. Subsequent investigations by US health authorities revealed that Showa Denko made two changes to its L-tryptophan manufacturing process in 1989 that allowed the contamination of the product: (a) the company began using a strain of *B. amyloliquefaciens* that had been genetically engineered to produce larger amounts of L-tryptophan and (b) the company reduced the amount of carbon used to filter out impurities from the final product to make the process cheaper. Studies have shown that the disease-causing molecule appears only during purification. In addition, cases of EMS have been linked to L-tryptophan produced by Showa Denko as early as 1983, long before the company used a

genetically engineered bacterium. The risk information vacuum for GE food allowed such alarmist and erroneous versions of events to take root and flourish.

Toward the end of 1996, the Natural Law Party mounted a cross-Canada book tour featuring Dr. John Fagan, an anti-GM activist and molecular biologist. This tour received extensive coverage, where the exaggerated or erroneous claims promulgated by Fagan and others were publicized. In particular, the statement was made that "40 people were killed and thousands were crippled by exposure to gene-tinkered food" (Graham 1996). There was no data to substantiate this claim, but the damage had been done. Many people began to doubt the safety of GM foods.

Powell and Leiss (1997) describe how a risk information vacuum arises when, over a long period of time, those who are conducting the evolving scientific research and assessments for high-profile risks make no special effort to communicate the results obtained from these studies regularly and effectively to the public. Instead, partial scientific information dribbles out here and there and is interpreted in apparently conflicting ways, all of which is mixed with people's fears. Accordingly, while education programs are not very effective, communication and transparency are still vital.

16.4. FEEDING FEAR: CASE STUDIES

Society as well as nature abhors a vacuum, and so this space is always filled with information from some source. While stories like those discussed earlier spark fear, other information must feed it over time. For example, events reported in the media become the substantial basis of the public opinion; or an interest group fills the vacuum with its own information and perspectives; or the intuitively based fears and concerns of individuals simply grow and spread until they become a substantial consensus in the arena of public opinion. Examples of these are discussed in the text that follows.

16.4.1. Pusztai's Potatoes

On August 10, 1998, Dr. Arpad Pusztai of the Rowett Research Institute in Aberdeen, Scotland, reported that after feeding five rats potatoes that were genetically engineered to contain one or two lectins, proteins that are known to be toxic to insects, they observed, over a 110-day period, that some of the rats manifested stunted growth and impaired immune systems. Dr. Pusztai reported the findings, not in a peer-reviewed scientific journal, but on the UK *World in Action* television program. After an internal review of the data, it emerged that not only had Dr. Pusztai ignored the conventional route of scientific peer review, but the experimental design lacked appropriate controls. Potatoes themselves are full of poisonous chemicals in quantities that vary depending on how they are grown, a phenomenon known as *somaclonal variation*, and must therefore be uniformly grown for any feeding trail to be informative. Moreover, rats do not like to subsist on raw potatoes, and their diets must be supplemented. By August 12, 1998, Dr. Pusztai was suspended and subsequently forced to retire.

The Pusztai affair, as it soon became known, spawned significant media coverage with numerous allegations. On February 12, 1999, a group of 20 international scientists released a letter supporting the work of Dr. Pusztai, and specifically charged that the process of genetic engineering itself, in particular the use of the 35*S* cauliflower mosaic virus promoter, was to blame. The 35*S* promoter is widely used in the genetic engineering of plants, to turn specific genes on and off. Because of this widespread use, regulators in Western countries already demand evidence that any 35*S* insertion is stable and well characterized. Other feeding experiments involving the 35*S* promoter have simply not found the problems described by Pusztai and supporters. Most importantly, though, the potatoes grown by Dr. Pusztai would never have been approved in Canada, the United States, or the United Kingdom. Subsequently, the UK Royal Society concluded that "Dr. Arpad Pusztai's widely publicized research into the effects of feeding rats GM potatoes appears to be flawed, and it would be unjustifiable to draw from it general conclusions about whether GM foods are harmful to human beings or not." The Pusztai affair is repeatedly cited as proof of harm from GE foods despite this declaration.

16.4.2. Monarch Butterfly Flap

On May 20, 1999 John Losey and colleagues from Cornell University published a brief letter in the scientific journal *Nature* (Losey et al. 1999) that drew intense national and international media coverage (PEW 2002). The report concerned a laboratory study in which the leaves of milkweed plants in a greenhouse were artificially dusted with pollen from conventional and genetically engineered Bt corn plants at levels approximating what the researchers thought occurred in nature. Bt corn has been genetically engineered to contain the protein from a common soil bacterium *Bacillus thuringiensis*. In this study, 3-day-old Monarch caterpillars were placed on the leaves and allowed to feed for 4 days. The researchers reported that 44% of the Monarch larvae fed leaves coated with Bt-pollen died. No caterpillar died that ate leaves dusted with regular corn pollen or the control leaves. Larvae feeding on the Bt-dusted leaves also ate much less and were less than half the size of larvae that fed on leaves with no pollen. No attempt was made, however, to compare the pollen coverage of the leaves in the lab to the coverage that might commonly exist in or near a cornfield.

The authors correctly recognized that the study was limited in applicability, and that field tests would be required to determine the significance of this small artificial-environment study. Upon publication, Dr. Losey was quoted as saying "We can't forget that Bt-corn and other transgenic crops have a huge potential for reducing pesticide use and increasing yields. This study is just the first step, we need to do more research and then objectively weigh the risks versus the benefits of this new technology" (http://www.news.cornell.edu/releases/May99/Butterflies.bpf.html).

Losey soon found that his results had transformed into mutant tales of killer corn and sacred butterflies. *The New York Times* led on the front-page with a story entitled "Altered Corn May Imperil Butterfly, Researchers Say" in which one researcher described monarchs as the "Bambi of the insect world" (Kaesuk Yoon 1999). Greenpeace demonstrators dressed in Monarch butterfly costumes and simultaneously drop dead at a prearranged time, usually for the convenience of television cameras: great street theater, but lousy public policy.

The Losey study could best be described as preliminary research, in which the larger conclusions were not well supported by the results. The sample size was extremely small, and so results were not statistically relevant. Another serious issue is the use of event "176" pollen, which is a transgenic Bt line that was rarely planted in the United States. The study did not report the Bt dosage on milkweed leaves, which was likely much higher than dosage that would be consumed by monarch caterpillars in the field. Toxicity results were not reproducible and larvae were force-fed Bt pollen on very small leaf sections. Therefore, these larvae could not avoid the pollen by choosing other food sources like they would in nature. Overall, the broad conclusions that Losey's study implied were not supported by the data. Shelton and Roush (1999) discussed this and other Bt pollen studies, which relied on insufficient evidence, yet fueled the power of rumor. Like other stories, despite its misleading, inaccurate conclusions and poor study design, this study still influenced public opinion regarding genetically engineered plants. It can be argued that the two case studies described already triggered the "dark ages" of the public perception of agricultural biotechnology that was fueled, in large part, by incomplete and speculative science combined with overly enthusiastic news media needing to "make" stories (Stewart and Littmann 2008).

16.5. HOW MANY BENEFITS ARE ENOUGH?

In October 2000, the US Environmental Protection Agency stated in a comprehensive report that corn, cotton, and potato crops genetically engineered to repel pests offered "significant benefits" to farmers and few risks, even for monarch butterflies, giving an overwhelming stamp of approval to the technology as a way to boost yields, reduce farm chemicals, and lessen groundwater contamination. The report found that in 1999 alone, US farmers reduced pesticide costs by more than $100 million (www.epa.gov/scipoly/sap/). Nonetheless, real benefits are often ignored as speculative sociological risks are incited.

The question "Do you want fish genes in tomatoes?" has been used repeatedly by Greenpeace and other activists in campaign literature and media accounts (McHughen 2000). Yet the actual experiment to transfer an antifreeze protein from cold-water flounder to enhance the cold tolerance of field tomatoes was attempted only once in 1991 and was unsuccessful (Hightower et al. 1991). Nevertheless, when asked which foods in the supermarket are GE, consumers consistently cite vegetables, such as tomatoes, and fruit (IFIC 2002). While this is due partly to the short availability of the Flavr Savr tomato, it also demonstrates how memorable such evocative messages are to the public.

These are only a smattering of the dozens of examples of information intended to alarm rather than inform. By the fall of 1999, this combination of scientific naivety, media hype, and allegations of corporate conspiracy characterized any and all public discussions of the role of genetically engineered foods. So Greenpeace and the Council of Canadians, two activist groups, hoping to build on the success in stigmatizing GE food in Europe (particularly in the United Kingdom) held a public demonstration in front of a Loblaws supermarket in an affluent area of downtown Toronto, a Canadian beachhead into the United States (Fig. 16.1). Typical of the statements was that of Jennifer Story, health protection campaigner for the Council of Canadians, who asserted that "Genetically engineered foods have not been proven safe for human health and the environment. As the largest grocery chain in Canada, Loblaws has the obligation to take the lead, and take genetically engineered food off the shelf."

When public concern mounted in the United Kingdom and Europe in response to activist tactics, the scientific community, political leaders, and opinion leaders were largely silent. Even if they had spoken out, the effects would have been marginalized by the fallout from the mad cow crisis in Britain. On March 20, 1996, the British government announced what many already knew—that consumption of products from cattle with bovine spongiform encephalopathy, or mad cow disease, was leading to a new variant form of Creutzfeld-Jacob disease (vCJD). This disease struck the young and was particularly gruesome, leading to the victim's inevitable death. Millions of animals were killed at a cost of billions of dollars in lost trade. To date around 229 people have been infected from vCJD internationally. Mad cow disease clearly represented modern agricultural practices as science out of control even though it likely originates from spontaneous generation of pathogenic prion proteins.

Unlike European farmers, North American farmers were eager to sample and adopt the newly available GE seeds, and were prepared to enter the public debate to retain and ensure access to those tools. As Thomas Jefferson wrote in a letter to William Charles Jarvis, dated September 28, 1820, "I know of no safe depository of the ultimate powers of society but the people themselves; and if we



Figure 16.1. Greenpeace demonstration in front of a Toronto grocery store (*Source*: photo by Doug Powell).

think them not enlightened enough to exercise their control with a wholesome discretion, the remedy is not to take it from them, but to inform their discretion" (NRC 1989, p. 14).

More recently, GM food labeling laws have been passed in a few US states such as Vermont, which were largely fueled by social media. In 2013, the "March Against Monsanto" Facebook page was launched to single out one company as a target for grass roots protest against GMOs. As of 2015, the page has over 800,000 "likes" and organizes marches in various cities to "take back our planet."

16.6. CONTINUING DEBATES

16.6.1. Process vs. Product

Genetic variability is required to enhance traits deemed desirable by humans. Geneticists can travel the world searching for plants, animals, or microorganisms that posses a trait of interest, such as increased productivity or disease resistance. Desirable variability can be selected over generations of breeding. Genetic engineering, using the tools of molecular biology, allows further sources of genetic variability to be introduced into a particular organism.

But there are other techniques for creating genetic variability between the black and white of traditional breeding and genetic engineering. Traditional breeding rearranges many more genes and is much less controlled than genetic engineering. Yet, in light of this fact, traditional plant breeding is not regulated or the target for public unrest. For example, since the 1940s mutagenesis breeding has been used to induce genetic variability, especially in cereals, by exposing seeds to doses of mutagens—compounds that induce mutations in DNA—such as ionizing radiation or mustard gas. The mutations are often many and random. The practice is still used today, as are other techniques. Because of the randomness and the mechanism, this technique could be inherently risky to some people by altering benign proteins to those that could be toxic or allergenic. In comparison, genetic engineering more precisely alters plant genomes to yield particular properties. Should the products from any of these processes be regulated, or is it the process of genetic engineering that is inherently risky?

Proponents and critics have sparred on this point since the advent of genetic engineering, but the scientific community and North American regulators have consistently maintained that it is the end-product (not the process) that should be regulated. Varieties of potatoes and celery produced through traditional breeding were later discovered to contain unacceptably high levels of natural compounds. The Canadian Novel Food Act (1999) encapsulated the view that the endproduct should undergo a safety assessment regardless of its production method. It was reaffirmed by an expert panel of the US National Academy of Sciences (National Research Council 2000).

16.6.2 Health Concerns

In 1994, the Flavr Savr tomato became the first whole, genetically engineered food to be approved by the US Food and Drug Administration (FDA) and, subsequently, Health Canada. Results of rodent feeding trials, submitted as part of the dataset that regulators reviewed, showed no difference between conventional and genetically engineered tomatoes. It also showed that rats do not like tomatoes, which may not be too much of a surprise.

The experiment highlighted one of the difficulties in assessing the safety of genetically engineered foods: equivalency. Any commercial concern wishing to sell a genetically engineered food, or indeed any new or novel food, must demonstrate substantial equivalence to the appropriate regulatory body through molecular, nutritional, and toxicological data. For example, the genetically engineered field corn grown in North America (and now elsewhere) contains a gene from the common soil bacterium *B. thuringiensis*, and is known as Bt corn. Regulators and several international scientific panels reasoned that because humans have been ingesting Bt without effect for decades (it is also widely used as an organic spray), the Bt toxins (in this case, specific to the European cornborer) are proteins, and because any toxin protein remaining after processing would be quickly digested in the human gut, Bt corn is safe.

In other words, the Bt corn was found to be substantially equivalent to traditional corn. If, on the other hand, substantial equivalence is more difficult to establish, then the identified differences or new characteristics would be the focus of further safety considerations. The more a novel food differs from its traditional counterpart, the more detailed the safety assessment must be. Future products of agricultural biotechnology, where complex plant pathways are exploited to produce more nutritious foods, may require a more elaborate safety assessment. In contrast, the genetically engineered foods available today are the result of relatively simple gene transfers, harnessing systems that are based in nature.

However, the attempt to improve any food can lead to unexpected consequences. For example, in one instance a human allergen was transferred from one crop to another. During the preliminary assessment process, the company immediately discontinued the experiment. For the critics of biotechnology, the experiment proved that allergens could be transferred; and therefore, untold risks lay in the manipulation of food structure. For supporters, the incident showed that the regulatory system worked. Indeed, molecular work in agricultural biotechnology has contributed significant knowledge to the database of food allergens.

16.6.3. Environmental Concerns

Biological systems are fluid and dynamic so that organisms can adapt to changing environmental conditions. Farmers have known for decades that when they overuse a particular agricultural tool, they create an evolutionary selection pressure. In many cases, this leads to resistance, rendering the tool ineffective. Genetic engineering tools are no different. In agriculture, weeds can significantly reduce yields. Farmers have a number of options for controlling weeds in a cost-effective manner, including the use of approved and registered herbicides, crop rotations, and most recently, genetic engineering. In particular, several herbicide-tolerant soybean and canola varieties are now available. This technology may allow producers to grow a bountiful crop with fewer chemicals.

One concern with herbicide-tolerant crops is that the genes responsible for such tolerance could move or transfer to neighboring weeds, thereby allowing such a weed to flourish as it becomes resistant to a particular herbicide (in which case the weed could still be controlled using other management practices such as tillage or alternative herbicides). Gene flow occurs through either pollen or viruses that can naturally infect one plant and then move on to another. The same concern about resistance applies to insect-resistant crops, such as Bt corn. That is why corn producers who grow genetically engineered Bt corn are, for example, required to devote 20% of their acreage to non-Bt varieties. The development of resistance and the transfer of genes from one plant to another are common phenomenon in agriculture. Therefore, we must begin to develop solutions to these issues to maintain environmental health and ecological diversity.

16.6.4. Consumer Choice

Consumer choice is a fundamental value for shoppers, irrespective of science. Foods in Canada and the United States are labeled on the basis of health and nutritional data, but there are a variety of other voluntary labeling systems based on religious preference (e.g., kosher and halal meats), growing preferences (e.g., organic), or nutritional preferences (e.g., low-fat and low-salt). A market for biotechnology-free foods, labeled as such, has emerged to meet consumer demand. However, many consumers will continue to base their food selections on taste, price, and nutritional content before other considerations. Labeling guidelines must accommodate all of these values.

16.7. BUSINESS AND CONTROL

The private sector has shaped public concern over agricultural biotechnology, particularly within multinational corporations such as Monsanto, Syngenta, DuPont, and Bayer. Such a concentration of expertise advances the research priorities of industrialized countries while potentially sacrificing

the public good. This desire of big business to maximize profits without attention to social consequences is well known. Big business' bad reputation has fueled the juxtaposition between safety and enterprise for many years.

This is a debate that predates transgenic plants, since food production has a long history of corporate involvement. On June 29, 1912, a prospectus for a new company, Synthetic Products Company Ltd., was launched in Britain following extensive newspaper advertisements. A global rubber shortage from 1907 to 1910 prompted European researchers to search for a synthetic source. Company backers believed this discovery was right around the corner. A group at the Pasteur Institute in France had discovered a bacterium that converted starch into a fuel oil rich in both amyl alcohol and butanol. When the process was scaled up to industrial quantities by British scientists, the fermentation was altered, producing butanol, which had just been recognized as a key component of synthetic rubber manufacture; and previously imported acetone, a valuable component of explosives. As recounted by Robert Bud in *The Uses of Life: A History of Biotechnology*, the work had enormous commercial potential, and the scientists "exploited the breakthrough to the hilt." The prospectus, which greatly exaggerated the scientific achievements, netted £75,000 despite stiff opposition from plantation rubber interests. Predictably, the process for converting starch from potatoes proved cumbersome, and the factory never realized the hopes expressed in the 1912 prospectus. But a pattern had been established: coupling scientific enthusiasm with the public's—at least the financial public—willingness to believe that would characterize efforts to profit from biology over the next century.

In a capitalist society, such involvement is to be expected. The challenge is to find a balance between private profit and public good in an open and democratic manner. Nonetheless, farmers, processors, distributors, and others in the farm-to-fork continuum are constantly striving to improve the safety, quality, and efficiency of the food supply. Genetic engineering is one tool that, with vigilance and oversight, can help achieve those goals.

16.8. CONCLUSIONS

After a decade of sometimes fierce public debate, what has been accomplished? Better oversight, changes in practices, shifting of entrenched attitudes? A little of all, but nothing of significance has been gained. A portion of the population remains polarized into for or against, yet overall we are slowly moving toward a public discussion of risks and benefits.

There might, however, be an intractable gap between opinions held by the public and scientists. In 2014, in collaboration with the American Association for the Advancement of Science (AAAS), the Pew Research Center held a sizeable survey about timely science and public policy issues. The goal was to compare positions held by the public, that is, non-scientists and AAAS members, that is, scientists (http://www.pewinternet.org/interactives/public-scientists-opinion-gap/). The two groups were very far apart on many issues, including evolution, climate change, and GM plants. The survey reported that 37% of US adults believe that GM foods are safe to eat, compared with 88% of scientists. This 51 percentage point gap was the largest of all topics surveyed.

Meanwhile, the World Health Organization estimates that up to 30% of all citizens of so-called developed Western countries will get sick from the food and water they consume each and every year; thousands will die. If the same energy and effort spent on GE foods could be harnessed to create a culture that values microbiologically safe food, there would far fewer sick people. In addition, there is also a general technology trickle-down effect. Technology is typically created in the developed world to be applied eventually in the developing world. Needless debates and fear mongering can slow down innovation, which, in turn, negatively affect the people in the developing world. Is it possible that protests in well-fed Europe have led to starvation in Africa? This is certainly food for thought as we consider about the future of plant biotechnology.

LIFE BOX 16.1. TONY CONNER

Tony Conner, Senior Scientist, New Zealand Institute for Crop & Food Research; Professorial Fellow, Bio-Protection & Ecology Division, Lincoln University, New Zealand.



Tony Conner with a transgenic potato plant. Courtesy of Tony Conner.

Towards designer plants. The first transgenic plants were developed in 1983 while I was studying toward my PhD in plant genetics at the University of California, Davis. At the time, my research involved somatic cell selection in Nicotiana plumbaginifolia as a model system. Upon graduation it was an obvious step to move toward developing transformation systems for crop plants. I was very fortunate to be offered a position back in my home country of New Zealand to establish a research program in applying the emerging tools in plant biotechnology to crop improvement.

It was an exceptionally exciting time to be involved in plant science. My research initially focused on potatoes, asparagus, and a few other vegetable and arable crops. In those early days it was rewarding to be associated with the first examples of Agrobacterium-mediated transformation of monocotyle-donous plants (asparagus) and some of the very first field tests on transgenic plants. Research advancements in plant molecular biology were rapidly gaining momentum, and this was matched by the development of molecular tools for analyzing genetic variation in plant populations and technologies for genetic engineering in a diverse range of plant species.

Integration of these new technologies into breeding programs of crops presented some important challenges. Often the elite material of plant breeders destined to become the future cultivars for the agricultural industries was more difficult to work with than other laboratory-based model systems. This was especially the case for developing transformation systems for gene transfer via genetic engineering.

However, public concerns about the deliberate release of transgenic crops into the agricultural environment quickly changed research agendas. Considerable effort was required to participate in the public debate on the merits and biosafety of transgenic crops and absorbed much of my time for about a decade. During this time my research efforts were directed more to investigating the environmental impacts and food safety of transgenic crops.

More recently my research focus changed to refining vectors systems for gene transfer to plants. This work has been motivated by the need to eliminate components of vectors that regulatorys (regulatory systems) find less acceptable.

This eventually led to our development of intragenic vector systems, which involve identifying functional equivalents of vector components from plant genomes and using these DNA sequences to assemble vectors for plant transformation.

Gene transfer using intragenic vectors allows the well-defined genetic improvement of plants without the introduction of foreign DNA. Biologically, the resulting plants are not transgenic, although the tools of molecular biology and plant transformation have been used in their development. The genetic make-up of the resulting plants is equivalent to a minor rearrangement of the endogenous DNA sequences within the species. This is

very similar to "micro-translocations" that can occur naturally in plant genomes or as a consequence of deliberate mutation breeding. For the transfer of genes from within the gene pools of crops, intragenic vectors may help to alleviate some of the public concerns over the deployment of GM crops in agriculture, especially ethical issues associated with the transfer of DNA sequences across wide taxonomic boundaries. Nowadays, my research is moving toward functional genomics of potato to better understand how important traits are controlled by specific genes and their alleles. I envisage this will lead to valuable sources of gene sequences for transfer to existing elite potato cultivars via intragenic vectors.

Early in my career I never considered it would be possible, in my lifetime, for science to generate the full genome sequence of a higher organism. Yet, within the next 5–10 years the annotated sequence, at least for the gene-rich regions, of the genomes of all major crops will be known. This will provide unprecedented opportunities for mining the germplasm collections of plant breeders for novel alleles that represent variant versions of genes with altered functions. The resulting novel DNA sequences can then be used for highly targeted genetic changes in crop plants by transformation of elite crop cultivars.

The next few decades are going to be exceptionally exciting for plant genetics as research moves toward the targeted design and development of genetically enhanced plants for sustainable production of high quality and healthy food. My career has been an exciting and fulfilling journey so far. But I often think: "What if I was thirty years younger?" What a tremendous career opportunity modern plant genetics would offer.

LIFE BOX 16.2. CHANNAPATNA S. PRAKASH

Channa S. Prakash, Dean, Tuskegee University; Winner of the 2015 Borlaug CAST Communication Award.



Channapatna Prakash. Courtesy of Channapatna Prakash.

I hail from an urban family of modest means in India. During the 1960s, India went through a difficult period as crops failed from drought, resulting in widespread food scarcity. The United States immediately sent huge shipments of wheat, saving millions of Indian lives. My childhood memory is thus peppered with images of wheat bags with USAID logo of "two clasping hands." Scientist Norman Borlaug visited India bringing magical seeds of dwarf-wheat, from his breeding program at CIMMYT, Mexico, which produced two to four times more yield than the prior wheat varieties. M. S. Swaminathan crossed these varieties into Indian wheat strains, and the "Green Revolution" was born. In the 1960s, India had barely managed to produce 8 million tons of wheat, and now produces 95 million tons.

As a child, I was fascinated by these developments in agriculture in India. Although a city boy, I was spending my summer holidays in the villages accompanying my grandfather who was in agribusiness. My

mother, a biology teacher, instilled a sense of awe and wonder of the natural world in me. I was more fascinated by the birds, butterflies, basil, and bougainvillea than toys. Of course, growing up with very limited means also meant that if one wanted a toy, then you had to make it yourself.

When most Indian youth then aspired to go to medical or engineering school, I chose to major in agriculture. A range of subjects like genetics, microbiology, entomology, pathology, horticulture, and agronomy taught at the University of Agricultural Sciences (UAS) equipped me with a broad platform. During the senior year, I followed the footsteps of Borlaug to choose plant breeding and genetics as my major. A charming and affable professor, G. Shivashankar took me under his wings and mentored me in the science of crop improvement. The vibrant scholarly culture fostered my growth, and I benefitted greatly from professors and graduate students at UAS.

When I began my master's thesis research on genetics of disease resistance in cowpea, I visited Rob Williams, a plant pathologist at the newly formed ICRISAT in Hyderabad, who mentored me by providing valuable guidance and a glimpse of an international science career. During grad school, I also ventured into science communication, writing popular articles in local newspapers and magazines, and producing radio documentaries on farming and science issues. I finished my master's degree in 18 months, and took up my first job at the Indian Institute of Science, working on a green energy project, under a legendary professor Amulya Reddy. I began my Ph.D. program in 1979 at UAS, when I also noticed that a few of my classmates began leaving for the United States for their graduate studies. Despite my aspiration in that direction, I could not afford to pay for the application fee to US universities or for entrance fee for the TOEFL and GRE exams. I had met my future wife Leela at that time and wanted to marry her and move away too far off land. Fortunately, I was offered an admission and scholarship to the Australian National University in Canberra (which had earlier rejected me but a kindly professor in Forestry wrote asking me if I would work on trees instead of field crops). I would have worked on unicorns at that moment to get to Australia!

Australia was a transformative experience. I pursued doctoral research on genetics of host–pathogen interaction in poplar leaf rust system and published a dozen papers from my thesis. The intellectual atmosphere at the topmost university in Australia was incredible, and CSIRO labs were just across the street. I never missed a seminar at both places. After completing my Ph.D. thesis, I came across a flyer announcing a meeting in June 1985 at Princeton University, USA, by Student Pugwash. I researched the new topic (genetic engineering and society!) to write an essay for the application, mostly questioning the new technology while visualizing many possible risk scenarios. To my delight, I was invited to Princeton (all-expense paid) to represent Australia! Spending a week at this prestigious Ivy League university in the company of some profound thinkers, scholars, students, and scientists discussing the value and risks of emerging biotechnology helped change my opinion on this technology and was a pivotal juncture in my career. The Princeton workshop broadened my mind and helped me appreciate the larger societal issues with technology such as ethics, economics, and history. Student Pugwash invited me to chair the next workshop at Stanford University in 1987.

While my research was in classical genetics so far, the Pugwash meetings motivated me to learn more about molecular biology. During my postdoc stint at the University of Kentucky, I also spent time sitting through genetic engineering lectures by gifted teachers (Joe Chappell, Art Hunt, and Laura Lacy). I befriended graduate students in agronomy to learn about tissue culture and plant transformation after hours (Prashanth Bhat and Wayne Parrott). I learnt lab techniques in DNA markers from Dave Wagner whose lab was next door in our Forestry Department. The University of Kentucky also featured seminars by outside experts, and I met many emerging giants in crop genetic engineering such as Roger Beachy, Maud Hinchee, Bob Goldberg, and Robb Fraley. UK rural sociologists Larry Busch and Bill Lacy were interested in larger GMO issues, and I used to interact with these people a lot.

As a new faculty at Tuskegee University in 1989, I began working on sweet potato and peanut, crops of choice to legendary scientist

George Washington Carver. I set up a biotech lab in the same building that Carver had earlier worked (Milbank Hall). As there was no start-up funds, I contacted many private companies like American Cyanamid and Abbott Labs asking for used equipment, which they donated. I also acquired surplus equipment from NIH. An extraordinary professor, JHM Henderson, allowed me access to his plant tissue culture lab at Tuskegee University, and his friendship is among my fondest memories of that period. Fortunately, my very first research grant proposal to USDA was funded. I went on to establish the largest biotech lab at any HBCU or 1890 University, training minority students and pursuing research on transgenic sweet potato and peanut, plus also genomic research on these two crops. My lab was always a busy and bustling place with graduate students, postdocs, international visiting scientists, an army of undergraduate students, and even high school teachers and students during the summers. We had weekly lab meetings and also journal clubs to discuss research papers of interest. We had many invited speakers delivering seminars while interacting with my students and scholars. It is among my most satisfying feelings that many of my students are in professionally successful positions. In 1996, I hosted one of the largest ever meetings on GM crops with the support of the NSF. Stalwart scientists such as Roger Beachy, Chris Somerville, Rob Horsch, Ganesh Kishore, Ananda Chakrabarty, Eugene Nester and Terry Medley were featured speakers, but I also provided young upcoming scientists (a certain postdoc from UGA Parrott Lab called C. Neal Stewart comes to mind) to speak on GMO issues, including biosafety, ethics and larger societal issues. The Association of Research Directors at 1890 Universities gave me their highest award in 2013.

Beyond my lab science, I was always communicating to others on GM crops and its safety: writing newspaper and magazine articles, speaking at the local Rotary Club or high school, and participating in Listserv discussions. When GM crops were first commercially planted in 1996, the topic slowly caught the attention of media and public, and I was gradually drawn into the "GMO debate" and became a science communicator. A US senator's office called me asking me if they could place my op-ed in the Congressional Record, and soon followed with an invitation to visit the senator himself. My speaking schedule slowly started turning busier. US State Department arranged for my visit as a "science diplomat" to several countries where I would typically lecture at various locations; meet with the local policy makers, regulators, academic, and industry experts; and often end with an elegant dinner hosted by the US ambassador. I have now delivered more than 500 public lectures across 75 countries and in such places as the United Nations, US congressional forum, Aspen Ideas Festival, World Food Prize, World Agricultural Forum, annual meetings of many scientific societies, and major campuses (Harvard, Stanford, Columbia, Purdue, UCLA, Berkeley). I was invited to speak at the Vatican three times, and recently had the great honor of meeting Pope Francis himself, along with Ingo Potrykus who asked the Pope to "bless" a sample of Golden Rice.

So, while I continue to teach plant genetics, plant biology, and biotechnology courses at Tuskegee, I also teach (along with Bob Goldberg of UCLA and John Harada of UC Davis) a course in "genetics and society" through interactive video that enables my Tuskegee students to learn from other experts. As you can see, my world unexpectedly expanded far beyond my wildest dreams. Therefore, my advice to young scientists reading this book is: follow your inner child. Think beyond the technicality of science and strive to reach out to others sharing the joy of science while helping them appreciate the nuances of innovation in advancing our society. Your world will also expand.

REFERENCES

Covello VT (1983): The perception of technological risks: a literature review. *Tech Forecasting Social Change* **23**:285–297.

Covello VT (1992): Trust and credibility in risk communication. *Health Environ Digest* **6**:1–5.

Crichton M (1990): *Jurassic Park*. Alfred A. Knopf, New York, **416**.

Durant J (1992): *Biotechnology in Public: A Review of Recent Research*. Science Museum, London.

Graham D (Nov 20, 1996): Altering food called "dangerous experiment." *Toronto Star*, **E1**.

Gregory R, Slovic P, Flynn J (1995): Risk perceptions, stigma, and health policy. *Health Place* **2**:213–220.

Groth E (1991): Communicating with consumers about food safety and risk issues. *Food Technol* **45**:248–253.

Gupta N, Fischer ARH, Frewer LJ (2012): Socio-psychological determinants of public acceptance of technologies: a review. *Public Underst Sci* **21**(7):782–795.

Hightower R, Penzes BC, Lund E, Dunsmuir P (1991): Expression of antifreeze proteins in transgenic plants. *Plant Mol Biol* **17**(5):1013–1021.

International Food Information Council (IFIC) (2002): U.S. Consumer Attitudes toward Food Biotechnology: More U.S. Consumers Expect Biotech Benefits. Wirthlin Group Quorum Survey, available at: http://www.foodinsight.org/Content/3843/2001_Food_Biotechnology_Consumer_Research_Summary.pdf (accessed on October 23, 2015).

Kaesuk Yoon C (May 20, 1999): Altered corn may imperil butterfly, researchers say. *New York Times*, A1.

Kelley J (1995): *Public Perceptions of Genetic Engineering: Australia 1994*. Australian Department of Industry, Science and Technology, Canberra. Retrieved September 10, 2007, available at: http://www.internationalsurvey.org/Kelley_1995_Pub_Percept_Genetic_Engineering2.pdf (accessed on October 23, 2015).

Key S, Ma JK, Drake PM (2008): Genetically modified plants and human health. *J R Soc Med.* **101**:290–298.

Losey JE, Raynor LS, Carter ME (1999): Transgenic pollen harms monarch larvae. *Nature* **399**:214.

McHughen A (2000): *Pandora's Picnic Basket: The Potential and Hazards of Genetically Modified Foods*. Oxford University Press, Oxford, pp 1–10.

PEW Initiative on Food Biotechnology (2002): Three Years Later: Genetically Engineered Corn and the Monarch Butterfly. Retrieved September 10, 2007, available at: http://www.pewtrusts.org/~/media/legacy/uploadedfiles/wwwpewtrustsorg/reports/food_and_biotechnology/vfbiotechmonarchpdf.pdf (accessed on October 23, 2015).

Powell DA, Leiss W (1997): *Mad Cows and Mother's Milk*. McGill-Queen's University Press, Montreal, pp 123–152.

Sagan C (September 10, 1989): Why We Need to Understand Science. *Parade Magazine*, pp. 6–10. Available at: http://news.google.com/newspapers?nid=1842&dat=19890910&id=9VMeAAAAIBAJ&sjid=EcgEAAAAIBAJ&pg=3074,1740126 (accessed on October 23, 2015).

Sandman PM (1987): Risk communication: Facing public outrage. *EPA J* **13**:21.

Shelton AM, Rousch RT (1999): False reports and the ears of men. *Nat Biotechnol* **17**:832.

Smith A, MacKinnon JB (2007): *The 100-Mile Diet: A Year of Local Eating*. Random House, Toronto.

Stewart CN Jr, Littmann M (2008): Media Treatment of Plant Biotechnology. Available at: http://agribiotech.info/more-details-on-specific-issues/files/StewartMediaMarch803.pdf (accessed on October 23, 2015).

US National Research Council (NRC) (1989): *Improving Risk Communication. Committee on Risk Perception and Communication*. National Academies Press, Washington, DC.

NRC (2000): *Genetically Modified Pest-Protected Plants: Science and Regulation*. National Academies Press, Washington, DC.

CHAPTER 17

The Future: Advanced Plant Biotechnology, Genome Editing, and Synthetic Biology

WUSHENG LIU AND C. NEAL STEWART, Jr.

Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee

17.0. CHAPTER SUMMARY AND OBJECTIVES

17.0.1. Summary

Even though plant biotechnology has been wildly successful and has literally transformed agriculture, an emerging engineering principles-driven plant synthetic biology is expected to play an important role in agriculture with the improvement of existing functions or the creation of new products in plants. Discussed here are the foundational principles, design cycle, components, and enabling tools of plant synthetic biology. Among the enabling tools, site-specific recombinases, ZFNs, TALENs, and clustered regularly interspaced short palindromic repeats (CRISPR) are expected to play a critical role in plant synthetic biology. A few examples of recent applications of what these new tools are making include synthetic elicitor- and soybean cyst nematode-inducible promoter construction, a device for monitoring auxin-induced plant indole-3-actic acid (IAA) degradation in yeast, and circuits for phytosensing of explosives or bacterial pathogens in transgenic tobacco and *Arabidopsis*.

17.0.2. Discussion Questions

1. What is plant synthetic biology, and what is the difference between synthetic biology and biotechnology?
2. What are the fundamental principles and design cycles of synthetic biology?
3. What are site-specific recombinases, ZFNs, TALENs, and CRISPR, and how might they alter the future of plant biotechnology?
4. Describe the definition of orthogonality.
5. Describe the design and construction of synthetic elicitor- and soybean cyst nematode-inducible promoters, a device for monitoring auxin-induced plant IAA degradation in yeast, and circuits for the phytosensing of explosives or bacterial pathogens in transgenic tobacco and *Arabidopsis*.
6. How do you think that plant synthetic biology will change agriculture and our daily lives in the future?

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

17.1. INTRODUCTION: THE BIRTH OF SYNTHETIC BIOLOGY

Today's world population of 6.5 billion humans is expected to reach 9.4 billion by the year 2050, which essentially requires a doubling of 1980s agricultural production. Will we be able to reach this lofty goal? Can agriculture continue to advance sufficiently to meet the world's projected needs such as feeding and clothing a rapidly growing population while maintaining environment and health quality in the next 40 years? To achieve this ultimate goal, we need new tools to produce new genetics to increase crop yield, stress resistance, and sustainability. There is also a need to genetically manipulate plants to perform novel functions and tasks such as phytosensing and the biosynthesis of new and valuable products. The emerging discipline of synthetic biology is expected to play a role not only by improving existing functions in plants but also in endowing new functions.

The term "synthetic biology" was first used by Barbara Hobom in 1980 to describe genetically engineered bacteria: living organisms that had been redesigned and modified (i.e., synthesized) (Hobom 1980). Later on, this term was reintroduced by Eric Kool and others in 2000 to describe the synthesis of artificial organic molecules that work in living systems (Rawls 2000). Then, the term was used to describe the efforts to "redesign life" (Szostak et al. 2001; Benner 2003), or to extract and characterize interchangeable parts from living systems that can be tested individually, then reassembled into living systems (Gibbs 2004).

Following the field's inaugural devices (i.e., the "genetic toggle switch" (Fig. 17.1a; Gardner et al. 2000) and the "repressilator" (Fig. 17.1b; Elowitz and Leibler 2000)), the principles of synthetic biology were first demonstrated using bacteria (Slusarczyk et al. 2012). The toggle switch consists of two constitutive promoters driving two different repressors, which, in turn, inhibit the transcription from the opposing promoters. When the toggle switch is implemented in *Escherichia coli*, it provides a synthetic, bi-stable gene regulatory network controlled by two inducers (Fig. 17.1a). To construct the repressilator in *E. coli*, three synthetic repressor–promoter pairs are arranged in a *cis* configuration. This allows expression of each repressor, which inhibits the transcription from the next promoter, which then provides an oscillating network that can be monitored by reporter gene expression from the last promoter (Fig. 17.1b). The result is a pulsating reporter gene.

This first wave of synthetic biology (Purnick and Weiss 2009) focused on proof-of-concept synthetic circuits described earlier as well as other "hey, look what synthetic biology can do" applications such as regulating the population of cells (You et al. 2004), logic-gates for whole-cell biosensors (Kobayashi et al. 2004), regulating the timing for fermentation processes (Ellis et al. 2009), and enabling bacteria to "see" light with the goal of using microbes to detect images (Levskaya et al. 2005, 2009; Tabor et al. 2009). Important real-life biomedical and industrial problems are being advanced, such as using bacterial cells to combat cancer (Anderson et al. 2006), increasing defense enhancement for bacteriophage (Lu and Collins 2007, 2009), and making new drugs (Ro et al. 2006). In addition, a huge milestone was reached in synthesizing and installing *de novo* genomes and sub-genomes. Along those lines, the phage $\Phi X174$ has been reconstructed via DNA synthesis (Smith et al. 2003), refactoring of the phage T7 (Chan et al. 2005), the creation of a bacterial cell harboring a synthesized 1.08 Mb *Mycoplasma* genome (Gibson et al. 2010), and the synthesis of the right arm of chromosome IX and a portion of chromosome VI in yeast (Dymond et al. 2011).

Although many of these advancements have been made within the realm of microbial synthetic biology, synthetic biology using plants is still is a long way behind that of bacteria (Liu and Stewart 2015). As the second wave of synthetic biology is creating systems-level circuitry (Purnick and Weiss 2009), we should be able to utilize the identical strategies to improve plants (Liu and Stewart 2015). Indeed, it has been demonstrated that some bacterial components can be introduced into the plant "chassis" using a synthetic biology approach for phytosensing of chemicals—see the following text for details (Antunes et al. 2009, 2011).

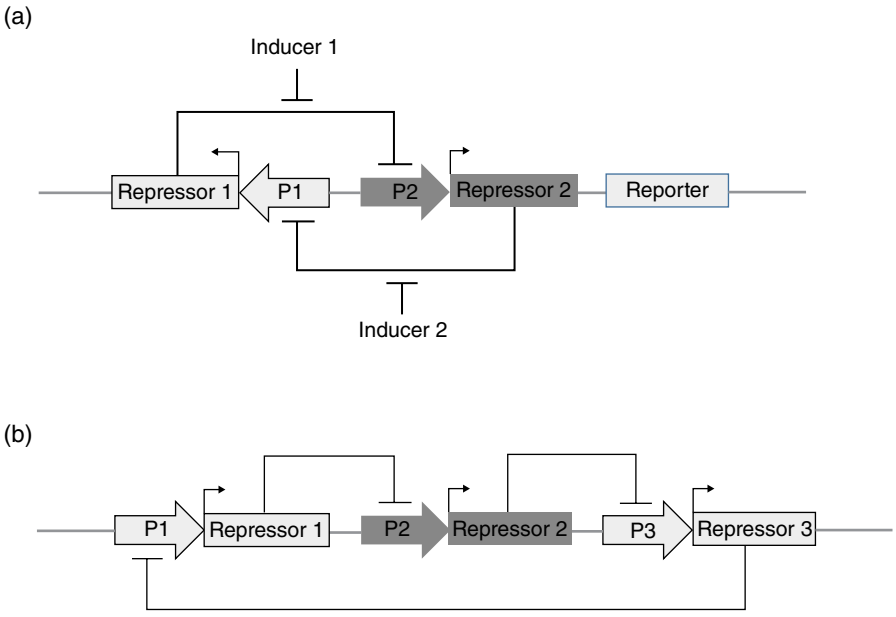


Figure 17.1. Two examples of synthetic circuits. (a) Genetic toggle switch: repressor 1 is induced by inducer 1 when driven by promoter 1 (P1), and inhibits transcription from promoter 2 (P2). Repressor 2 is induced by inducer 2 when driven by P2, and inhibits transcription from P1. Switching on the reporter gene expression can be achieved via transient application of either inducer (Gardner et al. 2000). (b) Repressilator: three synthetic repressor–promoter pairs are arranged in *cis* so that expression of each repressor can inhibit the transcription from the next promoter, which provides a periodically oscillating network can be monitored by reporter gene expression driven by the last promoter (*Source*: Adapted from Elowitz and Leibler (2000).)

17.2. DEFINING SYNTHETIC BIOLOGY FOR PLANTS

The capabilities of synthetic biology make a good fit for improving crop genetics. Synthetic biology aims to apply engineering principles to either the design and alteration of natural systems or the *de novo* construction of artificial biological systems that exhibit predictable behaviors in organisms (Schwille 2011). By utilizing well-characterized building blocks and mathematical modeling for rational design and synthesis of novel systems or functions, the development processes for plant synthetic biology are similar to working on a car. An older design can be modified for higher efficiency or a new model can be built from parts.

Synthetic biologists may view synthetic biology as being more engineering than biology, or vise versa. In complex organisms—in our case, plants—where there are still sizeable biological knowledge gaps, biology and engineering principles are likely equally important (Andrianantoandro et al. 2006). Synthetic biology can also be seen as a natural extension of plant biotechnology—it just uses newer tools, especially computational tools that were not available until recently. Biotechnology relies mainly on recombinant DNA technology using intact natural components. However, synthetic biology relies on synthetic components that are most often computationally designed and then synthesized in the lab.

17.2.1. Design Cycles of Synthetic Biology

Plant synthetic biologists should be able to set up their goals (i.e., desired features and functions of their favorite plant), design synthetic devices that meet those requirements, and implement their design into the plant chassis. Synthetic biologists, being part engineers, think of the plant body as a

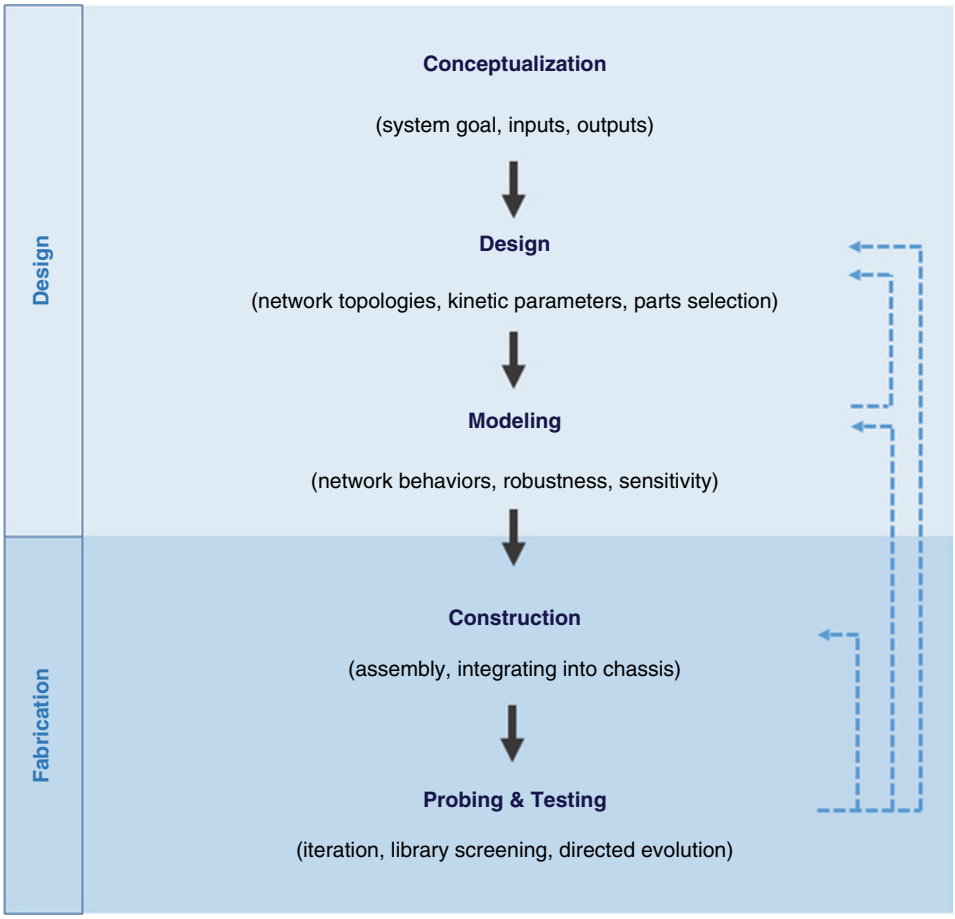


Figure 17.2. Design cycle for plant synthetic biology (reprinted with permission from Liu and Stewart 2015). An ideal design cycle consists of five stages: conceptualization, design, modeling, construction, and probing and testing. Conceptualization defines the overall goals of the synthetic project in response to inputs and outputs. Computer-aided design (CAD) assists in the optimization of network hierarchies, parameters and parts selection. Modeling helps analyze network behaviors, sensitivity and robustness. Construction is the assembly of synthetic devices and its loading into chassis. Optimization and multiple iterations (broken lines) are used to optimize the synthetic devices and systems.

chassis to bolt on new parts. The engineering/design cycle typically consists of the following five stages: conceptualization, design, modeling, construction, and probing and testing (Fig. 17.2; Alterovitz et al. 2009; Rollie et al. 2011; Slusarczyk et al. 2012).

Conceptualization defines the overall goals of the project in response to inputs and outputs so that genetic parts and assembly strategies are chosen to fulfill the goals. However, the prediction of macromolecule behavior is still an intricate task, as is their behaviors in complex plant systems.

Design and modeling, given the aforementioned uncertainties, heavily rely on computers. Computer-aided designs can be used to help determine and optimize network hierarchies, kinetic parameters, and parts selection.

Construction is the fabrication and assembly of synthetic circuits involved in standardized cloning and *de novo* DNA synthesis, and integration into plant chassis by either organelle or nuclear transformation. However, a complex system, especially with nonlinear interactions, may exceed our manageable abilities. It is challenging if not impossible to study complex system

behavior by studying the components separately. The synthetic biological approach is always challenged by variability, unpredictability, incompatibility, and not-so-well characterized components.

The trial-and-error approach combined with fine-tuning is always required to optimize the design cycle. Thus, optimization is used for experimental probing and testing. Moreover, multiple iterations (shown in Fig. 17.2 as broken lines) may be needed to obtain a series of approximate solutions and finally the desired functions and behaviors.

17.2.2. Foundations of Synthetic Biology

The most important foundational engineering principles for synthetic biology are decoupling, abstraction, modularization, and standardization (Endy 2005; Slusarczyk et al. 2012), which will be discussed later.

Decoupling and abstraction are related. Decoupling allows researchers to break down a complicated problem into many smaller problems that can be addressed individually so that the resulting work can be united together to produce a functional whole. For example, engineers can decouple a design from its fabrication. So decoupling focuses, ultimately, on parts design, subsystems design, and systems design, and then manufacturing of parts, and so on.

Abstraction separates hierarchies into workable levels. Then engineers can work at any level of complexity independently with limited and principled information exchanges between levels. For example, they can deconstruct an automobile into systems such as the body, drivetrain, electrical and electronics, interior, and miscellaneous parts. The systems can be deconstructed into subsystems, and each subsystem is deconstructed into auto parts that can be designed, constructed, and tested individually before being assembled to be a functional subsystem, then system, then automobile. A good example of an abstraction hierarchy that supports the engineering of new genetic systems is displayed in Figure 17.3. Utilizing the automobile analogy again, we can deconstruct and reconstruct the genetic system. Abstraction levels are listed as DNA, parts, devices, and systems. DNA is genetic material, which can be assembled rationally to form "parts" with basic biological functions, for example, genes and promoters. "Devices" are constructed with any rational combination of different "parts" to perform human-defined functions, such as genetic constructs and circuits, which can also be assembled or integrated into a chassis to generate an artificial genetic "system" for any desired purpose.

Modularization is used to define a functional unit with intrinsic properties independent of its outer connections, and it plays a fundamental role in the prediction of the system's behavior based on the behavior of its components. Synthetic biologists can achieve modularity by minimizing the effects of a component on the same or upper hierarchical level(s), or the effects of a downstream process on an upstream process (known as retroactivity). Examples of modules are groups of coexpressed genes, MAPK cascades to result in signal transduction in cells, and the machinery for protein synthesis or DNA replication.

Standardization is used to define and characterize interchangeable orthogonal biological building blocks as well as the standardized conditions for testing and improving each individual part. Orthogonality means that each added building block should not crosstalk with those present in the devices or systems or the plant chassis itself. Therefore, standardization relies on the ability insulate new parts from other parts as needed. Standardization relies on modularity and orthogonality of components, and allows components to be easily combined to form a complex system. Synthetic biologists are focusing on creating basic part libraries whose parts can be easily assembled for proper functions. For example, "BioBricks" have been identified and assembled (http://partsregistry.org/). These parts use a standard DNA cloning mechanism for their assembly and use in bacteria. The simplification of a technical process by decoupling, abstraction, modularization, and standardization of parts can significantly improve out ability to "do" synthetic biology. We have to make the components (parts).

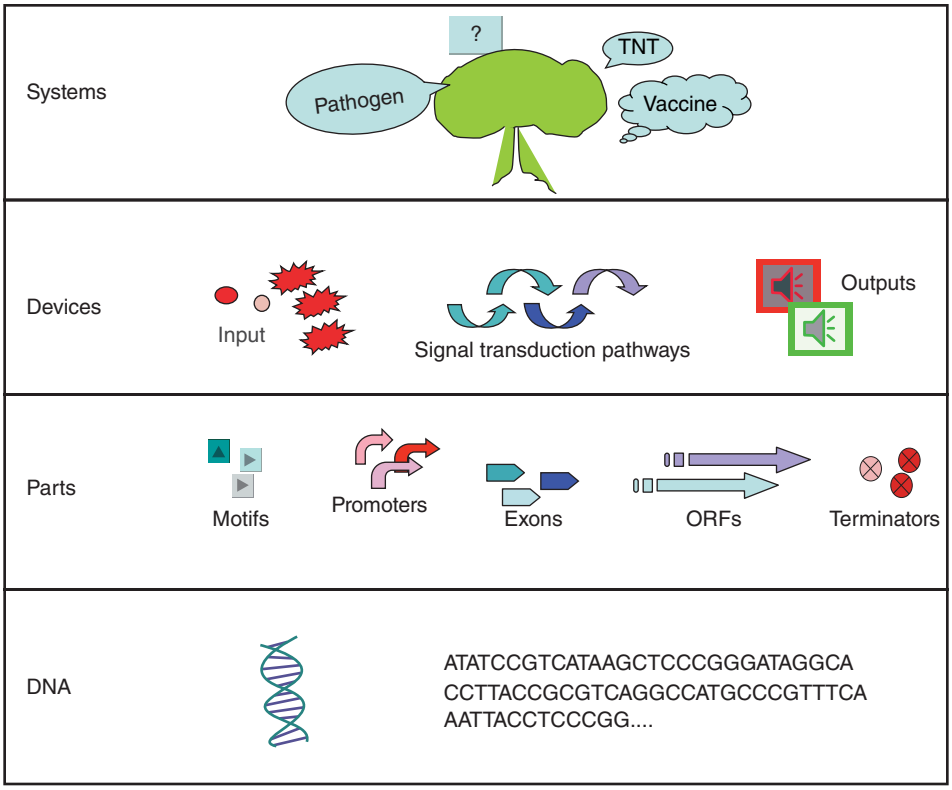


Figure 17.3. An abstraction hierarchy that supports genetic engineering of biological systems. The parts (such as promoters, ribosome-binding sites (RBS) in prokaryotes, coding regions, and terminators) are DNA sequence-based and are sometimes context-dependent, but can be engineered rationally to produce different devices such as inputs, logic gates and outputs, which permit assembly into artificial systems for further practical, desirable applications.

17.2.3. Components of Plant Synthetic Biology

Plant synthetic biology components (or parts) are made to affect the host organism and be analyzed at various levels, including genome, transcriptome, proteome, and metabolome. Mostly, parts are made of DNA whose sequence is typically computer-designed. The parts could be *cis*-regulatory elements (or motifs), promoters, transcription initiation sites, exons, protein domains, protein-coding open reading frames (ORFs), or terminators. Motifs can be standardized and used for synthetic promoter engineering. Exons and protein-coding domains can be used to generate synthetic chimeric genes and proteins. Ideally, these components can be tailored and organized in a highly modularized fashion so that they can be assembled modularly to produce different hierarchical functions, such as new synthetic genes, pathways, chromosomes, genomes, and/or conglomerate biological devices and networks.

Biotechnology has typically used a more natural version of the ones mentioned earlier—recombining promoters, genes, and so on, found in nature somewhere. The replacement of natural components with synthetic parts can significantly reduce the context dependency and "noise," while enhancing the functionality of the components or their interaction with other parts. This expectation led to the development and inventory of the BioBrick assembly standard (http://parts.igem.org/Assembly:Standard_assembly) and the Registry of Standard Genetic Parts (http://parts.igem.org/Main_Page), a database of BioBrick parts. The genetic parts cataloged in these databases allows for

easier assembly into larger genetic devices with predictable functions. Some (if not all) of these BioBricks could also be used in plant synthetic biology. Moreover, their plant subset, Phytobricks (www.plantfab.org), has also been established.

17.3. ENABLING TOOLS FOR PLANT SYNTHETIC BIOLOGY

There are several specific enabling tools either specific for plants or useful in plants that will be discussed. These include tools for computer-aided design (CAD), the production of synthetic promoters, DNA assembly and synthesis, loading into the plant chassis, and precise genome editing. For DNA assembly and synthesis and loading into plant chassis, please see Chapters 7 and 10. They are really carryovers from biotechnology and are being extended into the realm of synthetic biology. In this chapter, we will discuss CAD, synthetic promoters, and precise genome editing. This last topic is exploding in usefulness in all biology, including plants.

17.3.1. Computer-Aided Design

One of the most important aspects of synthetic biology is the development of CAD tools for design and modeling. Plant synthetic biology researchers can use CAD tools to specify outputs, analyze network topologies, construct systems and implement, and evaluate the design *in silico*. Many of such computational tools have been generated for simpler prokaryotic systems (see more details in Liu and Stewart 2015), and novel tools are being developed for plant synthetic biology. For example, *CellModeller* developed by Jim Haseloff's lab (http://www.haseloff-lab.org/) in Cambridge, UK, can be used to model cell–cell interactions during plant morphogenesis and to perform synthetic plant design.

17.3.2. Synthetic Promoters

Synthetic promoters were introduced in Chapter 10. Synthetic promoters can serve as the key regulatory components for regulation of gene expression in designed circuits. Just like their endogenous counterparts, synthetic promoters can be constitutive, inducible, spatial (tissue-specific), or temporal (developmental). They are designed to be so by consisting of a limited number of defined motifs, which act as binding sites of key transcription factors. Thus, systems biology (see Chapter 6) is important to understand the "wiring diagrams" and cross talk between transcription factors and the promoter motifs on which they bind. Known plant motifs can be easily detected when compared to the three well-known plant *cis*-motif databases: PLACE, PlantCARE, and TRANSFAC. However, novel motif discovery has to rely on bioinformatics-based *de novo* motif discovery. Thereafter, experimental approaches can be applied to further test the detected motif function, using 5′-end serial deletion (deconstruction) or addition of motifs (reconstruction) individually or in combination. In addition, the construction of combinatorial promoter libraries may offer great value in plant synthetic promoter engineering.

While it is arguably the most advanced class of parts for plant synthetic biology, the list of plant synthetic promoters is short (see Chapter 10). So far, even with a very broad definition of synthetic promoter, that is, "synthetic sequence" instead of naturally existing sequences in plants, only a limited number of synthetic (nonnatural) promoters have been created and used for constitutive and/or bi-directional gene expression in plants, as well as those that are inducible and tissue-specific.

17.3.3. Precise Genome Editing

Genome editing means the generation of desired genome modifications (i.e., gene disruption, addition, or correction) through the production of a double-strand break (DSB) in a specific genome site and subsequent DNA repair. In plants, DSBs are repaired mainly by error-prone nonhomologous

end-joining (NHEJ), or less frequently, by homologous recombination (HR). DNA repair is often imperfect and results in mutations in the targeted genomic sites. There are several genome editing tools that have been used in plants, which will be discussed later—from oldest to newest technologies: site-specific recombinases, zinc finger nucleases (ZFNs), transcription activator-like effector (TALE), TALE nucleases (TALENs), as well as CRISPR.

17.3.3.1. Site-specific recombinases. Bacterial viruses, known as phages, often integrate their DNA into the genome of a bacterial host at a designated site. This site-specific integration process is highly efficient and depends on recombinase protein(s) encoded by the phage genome. Integration into the host genome enables a phage to hide within the bacterial genome for many generations until it elects to leave by a reversal of the integration process, or site-specific excision, which is also mediated by a phage-encoded excisionase protein(s). These site-specific recombinases can be used to break DNA in two directly orientated recognition sites. Although these recombination systems originate from prokaryotes or lower eukaryotes, many of them also function in higher eukaryotic cells.

The types of site-specific recombination systems that function in higher eukaryotes can be classified into several groups according to biochemistry. Some tyrosine recombinases, such as Cre, FLP, and R, have been widely used for transgenic removal with the help of a catalytic tyrosine for cleavage; they are bidirectional and fully reversible with *loxP, FRT*, and *RS* being their respective DNA recognition sites (Fig. 17.4a). The recognition sequences must be introduced into the target DNA—then the recombinase cuts at these sites and removes flanking DNA, but are fully reversible. These tyrosine recombinases do not need host-specific factors to function in plants.

Moreover, some serine recombinases—CinH, ParA, Bxb1, and PhiC31—confer irreversible excision in the absence of their helper protein excisionase in plants. CinH and ParA recombinases (also known as *resolvases*) use *RS2* and *MRS* as their respective recognition sites, whereas Bxb1 and PhiC31 (also known as *integrase*) use *attB* and *attP* (Fig. 17.4b and c), and yield hybrid product sites *attL* and *attR* after excision. Recall that some of these are powerful vector-production tools (see Chapter 7). These recognition sites are much longer than those tyrosine recombinases, which greatly decreases the possibility of off-target effects.

To remove the marker gene (or a transgene), transgenic plants containing recognition sites flanking a marker gene (or a transgene) can be crossed with plants harboring their respective recombinases; they can also be re-transformed with their respective recombinases (Fig. 17.4d). Alternatively, recombinases can be controlled by an inducible or tissue-specific promoter in a construct containing the marker gene (or transgene) that is flanked by their specific recognition sites (Fig. 17.4e). Recently, these site-specific recombination systems have been used for transgene removal in several plants species and for other purposes (Wang et al. 2011). The efficiency of marker gene or transgene removal depends on excision efficacy, the uniqueness of the binding sites, and the binding specificity of the recombinases.

17.3.3.2. Zinc Finger Nucleases. Both ZFNs and TALENs are customizable protein-based systems to bind and cut a DNA sequence specific to a host sequence. They both use protein systems found in nature that bind DNA at specific sequences, and then fuse the protein to a nuclease. In both cases, the efficiency and precision of targeting depends on the DNA-binding domains that can be customized using synthetic biology, for example, computational design of the customized biding.

Zinc finger proteins (ZFPs) are one such tool used to exploit DBPs. ZFPs bind to DNA targets as a monomer consisting of a tandem array of typically three-to-six Cys2–His2 fingers and have 9–18 bp of specific target sequences. The engineered ZFPs have been widely used for gene activation or repression in plants when fused to either transcription activation or repression domains. Fusion of ZFs with a non-specific DNA cleavage domain of the *Fok*I endonuclease makes a novel nuclease, zinc finger nuclease (ZFN; Fig. 17.5a). After binding to the target site via the ZF domains, the *Fok*I endonucleases must dimerize to make DSBs within the spacer regions between two binding sites.

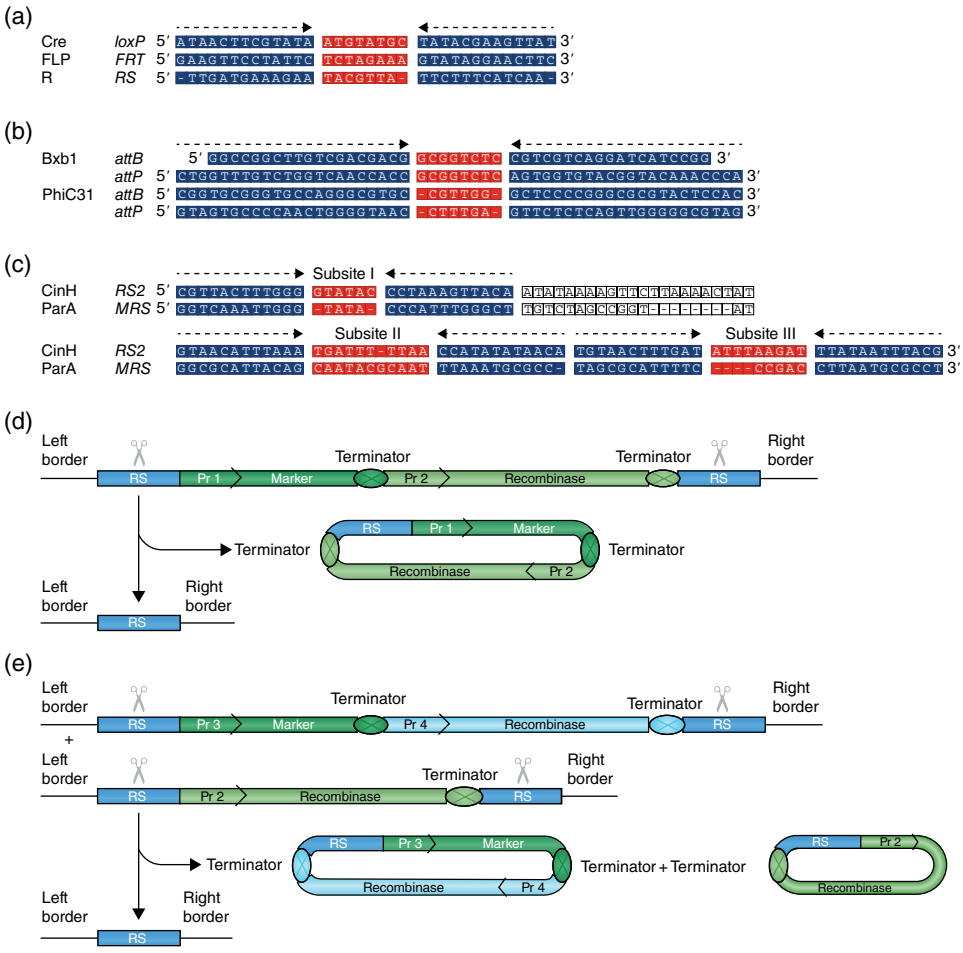


Figure 17.4. Site-specific recombinases (reprinted with permission from Liu et al. 2013b). (a) The recognition sites of tyrosine recombinases Cre, FLP, and R are *lox, FRT*, and *RS*, respectively. Each site contains two oppositely oriented, almost identical, repeats that allow reversible recombination. (b) The recognition sites of serine recombinases Bxb1 and PhiC31 are *attB* and *attP*. Both sites differ in sequence and yield hybrid product sites *attL* and *attR* after excision. (c) The recognition sites of CinH and ParA are *RS2* and *MRS*, respectively, which consist of three subsites: I, II, and III. Subsite I is the recombination site, while subsites II and III are accessory sites. The spacer sequences are shown in dotted arrows. (d, e) The recombinases can be used for marker gene removal from plant genomes. The recombinases can be inserted into the same construct as marker gene (d), or in another construct (e). The core nucleotides for strand exchanges in each recognition site are shown in red in panels a–c except the subsites II and III in panel c. RS, recognition site; Pr, promoter; LB, left border; RB, right border; ⊗, terminator.

The spacer length for ZFNs can be 5–7 bp in length, while the amino acid sequence of ZFN linkers of up to 20 amino acids in length could possibly affect the affinity for target sites. The specificity of ZFNs highly depends on (a) engineered DNA-binding codes, (b) numbers of the binding repeats, (c) spacer length, and (d) linker sequences between the binding and nuclease domains. The engineered nucleases have been utilized for gene disruption and mutational insertion or deletion or substitution by NHEJ, and for gene targeting, correction and even gene disruption via targeted gene addition by HR with the help of donor plasmids. The efficiency of site-directed mutagenesis is typically below 5%, but ZFNs have been exploited for a number of applications; thus, their utility has been proven (Cai et al. 2009; Ainley et al. 2013).

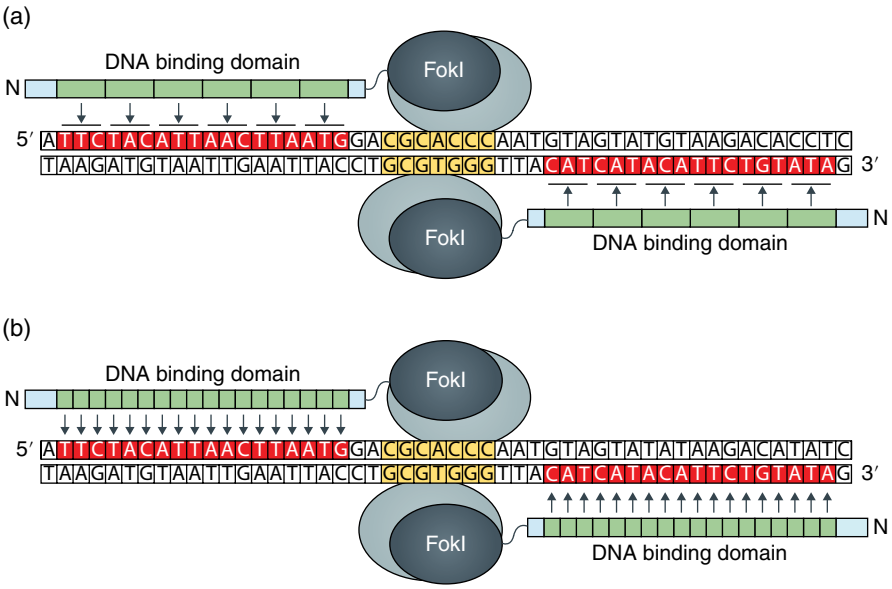


Figure 17.5. Engineered zinc finger nucleases (ZFNs; (a) and transcription activator-like effector nucleases (TALENs); (b) for targeted genome modification (reprinted with permission from Liu et al. 2013b). Each nuclease contains a custom-designed DNA-binding domain and the nonspecific DNA-cleavage domain of the *Fok*I endonuclease which has to dimerize for DNA cleavage within the spacer regions between the two binding sites. The spacer regions between the monomers of both nucleases are 5–7 bp and 6–40 bp in length, respectively. (*See insert for color representation of the figure.*)

17.3.3.3. TALENs. Another novel genome editing approach is using TALENs for double-strand break generation at any specific genomic location, providing favorable qualities for genome editing.

The more recently discovered TALEs are avirulence factors secreted by the pathogenic bacterium *Xanthomonas* that naturally bind to the promoter regions of plant resistance (R) genes and activate their expression. TALEs contain three domains, that is, an N-terminal translocation domain, a central DNA-binding domain, and a C-terminal domain with activation domain. The DNA-binding domain of *Xanthomonas* TALEs are composed of mostly 15.5–19.5 tandem, nearly identical repeats, with each repeat being 34 amino acids in length. The specificity of individual repeats is encoded in a *r*epeat-*v*ariable *d*iresidue (RVD) at positions 12 and 13 in each repeat. The four most common diresidues being NI, NG, HD, and NN on these two positions specifically bind to nucleotides A, T, C, and G, respectively. Thus, the code of designing TALEs to bind to any DNA sequence has been solved. Replacement of the activation domain of TALE with the nonspecific DNA-cleavage domain of the *Fok*I endonuclease generates a novel nuclease, TALEN (Fig. 17.5b). Like ZFNs, a pair of TALEN monomers can be designed to bind to two adjacent genome sites with the spacer regions being 6–40 bp in length, so that the *Fok*I nuclease can dimerize to cut DNA at a specific location.

ZFNs or TALENs can be used for targeted generation of resistance mutations, specific and permanent removal of undesired genes or selection markers in transgenic plants (reviewed in Mahfouz et al. 2014). In addition, they could also be used for gene stacking, gene replacement, and recovery of double mutants for tightly linked genes. In the presence of a donor vector, they could also be used for targeted genomic integration of synthetic circuits into plant hosts in the near future. However, TALEN repeat units are context-independent and thus provide more predictable sequence specificity than ZFNs whose desired specificity has to be screened by complex expression libraries (reviewed in Liu et al. 2013b).

17.3.3.4. CRISPR. CRISPR is the newest genome editing tool. Unlike ZFNs and TALENs, it uses a single DNA nuclease called Cas9 and a RNA-guided mechanism to target DNA for cutting. Like the other tools, the exact location for cutting is designer-friendly and much easier to deploy than any other tool-to-date. Thus, we supply a lot of background information about this powerful system that will likely transform how plants are transformed.

CRISPR were first identified in *E. coli* in 1987 as intergenic loci containing multiple short direct palindromic repetitive elements of 29 nucleotide length interspaced with 32-nucleotide spacer sequences (Ishino et al. 1987). It was also found that an upstream (A+T)-rich leader sequence serves as a promoter element for CRISPR transcription. Subsequently, bioinformatic analyses have shown that CRISPR loci are flanked by a few CRISPR-associated (*Cas*) genes. Some *Cas* genes function as helicases or exonucleases. Following the gradual discovery of this unique repeat–spacer–repeat pattern with different unit numbers in some bacteria and archaea (but not in eukaryotic chromosomes), it was found that the repeat sequences within a CRISPR locus are conserved but vary in both sequence and length in different loci. The spacers were demonstrated to show sequence homology to some viruses and plasmids in 2005, leading to the hypothesis that CRISPR could play a role in immunity against invading viruses and plasmids (Fig. 17.6; Bolotin et al. 2005; Mojica et al. 2005; Pourcel et al. 2005). Thus, bacteria and archaea containing CRISPR loci respond to viral or plasmid attacks by preferentially integrating short sequence fragments (i.e., protospacers) of invading DNA into their own genome as spacers at the leader end of the CRISPR loci. Transcription of the repeat–spacer elements can produce precursor crRNAs (pre-crRNAs), which are subsequently truncated to be short CRISPR RNAs (crRNA). The crRNA can direct Cas proteins for subsequent attack on its matching protospacer sequences of intruders. Therefore, the spacers are derived from incorporated invading genetic elements, and provide a genetic memory of infection.

There are three types of CRISPR/Cas systems, with type II system being best studied. The protospacer sequences in the type II system always immediately precedes a NGG motif, which is called protospacer adjacent motif (PAM); thus the protospacer sequences were not randomly selected from invading DNA. Type II system consists of four *Cas* genes, one of which is always endonuclease *Cas9*. Cas9 together with RNase III are believed to be responsible for crRNA processing and the silencing of invading DNA, with the help of a transactivating crRNA (tracrRNA). The tracrRNA is found to be coded upstream of the CRISPR locus on the opposite strand and contains a 25-nucleotide sequence that is complementary to the repeat sequences in the pre-crRNA. Fusion of tracrRNA with pre-crRNA (i.e., single guide RNA or sgRNA) has been demonstrated to be efficient for sequence-specific cleavage of target DNA (Jinek et al. 2012). The sgRNA associates with Cas9 and forms a CRISPR-associated ribonucleoprotein complex that recognizes invading DNA via base-pairing interactions between the sgRNA spacer sequence and a complementary sequence on either the coding or noncoding strand of the foreign DNA.

So far, this system has been used to target various genes in plants to mainly knock out genes. It has also been used in gene mutagenesis to introduce specific sequences in plants. The targeted mutagenesis rate in the earliest studies was about 2–5% using agroinfiltration and 6–39% using protoplast assays (Li et al. 2013; Nekrasov et al. 2013; Shan et al. 2013). It is clear that the CRISPR/Cas9 system is a "game-changer" in plant biotechnology and will enable both commercial development as well as basic research (Belhaj et al. 2015).

17.4. SYNTHETIC BIOLOGY APPLICATIONS IN PLANTS

The first fruit of plant synthetic biology research is ripening. So far, there are only a few good examples published in plant synthetic biology, such as synthetic elicitor- and soybean cyst nematode-inducible promoter construction, a device for monitoring auxin-induced plant IAA degradation in yeast, and circuits for the phytosensing of explosives or bacterial pathogens in transgenic tobacco and *Arabidopsis*.

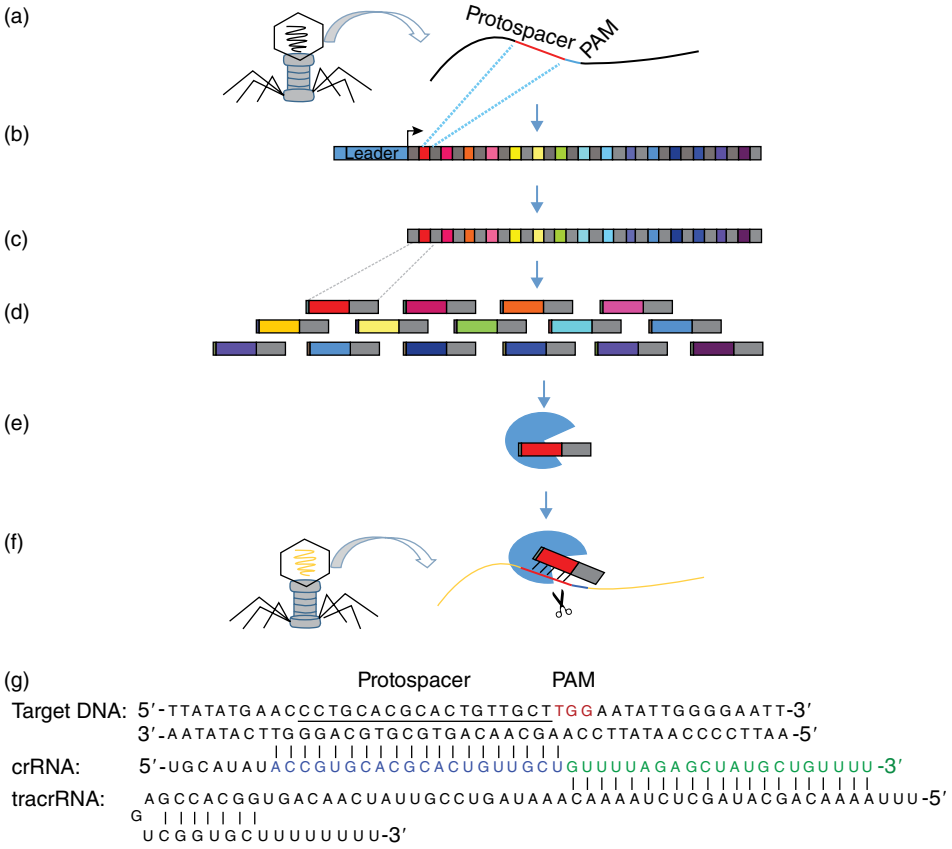


Figure 17.6. Outline of the CRISPR-Cas defense pathway (modified from Terns and Terns (2011) and Jinek et al. (2012)). (a) A short viral or plasmid DNA sequence (protospacer) upstream (for type II system) to the protospacer adjacent motif (PAM) is acquired and integrated into the host CRISPR locus adjacent to the leader sequence. (b) The CRISPR locus consists of invader-derived spacers with similar sizes (multiple colors) interspersed with short direct repeats (dark gray) and the leader sequence. (c) The transcription of the CRISPR locus using the leader as promoter produces pre-crRNA. (d) The pre-crRNA is processed to be mature crRNA, which typically contains an 8-necleotide repeat sequence at the 5′-end and a 20-nucleotide repeat sequence at the 3′-end. (e) Each crRNA binds to Cas9 protein (blue) with the ability to target different protospacers. (f) The binding of the crRNA-Cas9 ribonucleoprotein complex to a target sequence in the same or phylogentically closely related invader genome through base pairing leads to a double strand break in the target site. (g) Illustration of base pairing between crRNA and a target site on the foreign DNA, and between crRNA and tracrRNA in the type II CRISPR-Cas9 system. (*See insert for color representation of the figure.*)

17.4.1. Synthetic Inducible Promoters

Plant synthetic promoter engineering is highly limited by the availability of well-characterized motifs and computational modeling software. Current synthetic promoter construction mainly depends on reconstructive analysis that combines known motifs together or inserts known motifs into well-known core promoters for an enhanced (conditional) expression ability. The combination of a bioinformatics-based *de novo* motif discovery with experimental functional analysis is expected to play a key role in plant synthetic promoter engineering.

A pioneering example is that of Koschmann et al. (2012) who combined five *de novo* motif discovery tools from the binding-site estimation (BEST) suite with experimental analysis for novel elicitor-responsive motif discovery in *Arabidopsis*. *Arabidopsis* genes were selected for *de novo* motif discovery that were at least twofold up-regulated by one to six pathogen-related stimuli in the

PathoPlant database (http://www.pathoplant.de), which contains microarray experiments studying plant–pathogen interactions. The consensus sequences of the discovered motifs were used for functional analysis as well as synthetic elicitor-inducible promoter construction in *Arabidopsis*.

Liu et al. (2014) also used a set of seven *de novo* motif discovery tools for novel soybean cyst nematode (SCN; an important pest in soybean)-inducible motif discovery for use in soybean. They conducted *de novo* motif discovery among the promoter regions of 18 co-regulated genes that were selected from six published microarray studies on the compatible soybean–SCN interaction. The overlapping motif regions that were detected computationally by at least four out of seven bioinformatic tools were experimentally downselected using a stable transgenic soybean hairy root system, leading to the discovery of tens of core motifs of 5–7 bp in length. These newly discovered core motifs were then used for synthetic SCN-inducible promoter engineering.

Taken together, these two examples demonstrate the effectiveness of the combination of *de novo* motif discovery with experimental functional analysis for plant synthetic promoter engineering.

17.4.2. A Device for Monitoring Auxin-Induced Plant IAA Degradation in Yeast

Auxin plays a critical role in the regulation of plant growth and development. It interacts with its receptors F-box proteins TIR1/AFBs in order to activate gene expression by inducing the turnover of the auxin IAA. To comprehensively study the dynamic turnover of the auxin-induced plant IAA protein, Havens et al. (2012) overexpressed TIR1/AFBs and the yellow fluorescent protein (YEP)-IAA in different yeast (*Saccharomyces cerevisiae*) strains (Fig. 17.7). Without the presence of auxin or its receptors, the YEP-IAA fusion protein was stable in yeast. When yeast expressing either TIR1/AFBs or a YEP-IAA was mated and auxin was applied, the binding of auxin to its receptors and subunit II of IAA led to the ubiquitination and degradation of IAA. Thus, the dynamic turnover of IAA could be observed via YFP fluorescence measurement. This unique synthetic device permits precise control of input (i.e., external auxin application) and dynamic measurement of output (YFP fluorescence).

17.4.3. Circuits for Phytosensing of Explosives or Bacterial Pathogens in Transgenic Plants

Plants can detect and respond to various environmental and biological stresses. Making plants to sense and report in the environment has been greatly enabled by synthetic biology. Synthetic promoter construction using *cis*-regulatory elements and signal transduction pathways to create phytosensors for novel sensory functions is one of the few examples of an application in plant biology where synthetic biology approaches have already proven to be effective.

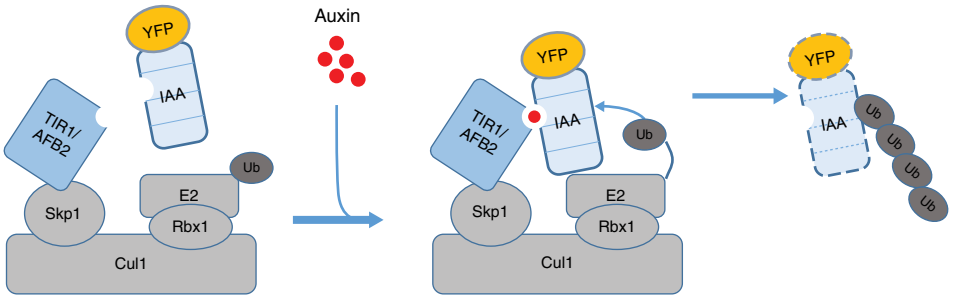


Figure 17.7. Device for monitoring auxin-induced plant IAA degradation in yeast *Saccharomyces cerevisiae*. The mating of yeast strains expressing TIR1/AFB2 or YFP-IAA and the application of auxin lead to ubiquitination and degradation of YFP-IAA. (*Source*: Modified from Havens et al. (2012).)

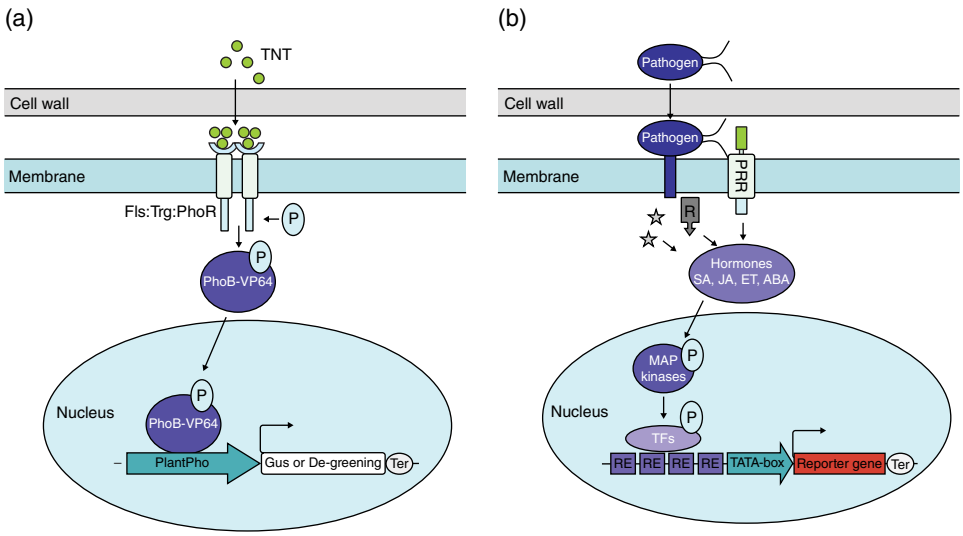


Figure 17.8. Circuits for phytosensing of TNT (a) or bacterial pathogens (b) in transgenic tobacco and *Arabidopsis*. (a) The TNT-inducible circuit contains a bacterial response regulator PhoB. Upon TNT contamination, the TNT-bound, computer-designed TNT receptors induce phosphorylation (P) of Fls:Trg:PhoR fusion protein which is cell membrane-localized and, in turn, induces phosphorylation of PhoB-VP64. The phosphorylated PhoB-VP64 moves into the plant nucleus and activates Gus expression or a de-greening system driven by a synthetic promoter (*PlantPho*). (b) The bacterial pathogen-inducible promoters contain four head-to-tail copies of hormone signal inducible *cis*-regulatory elements (RE; i.e., 4 × PR1, 4 × SARE, 4 × ERE and 4 × JAR). Pathogen attack can lead to activation of hormone defense pathways in transgenic plants via pathogen-secreted effectors (shown as stars), pattern-recognition receptors (PRRs), or acquired resistance (R) proteins. Then the MAP kinase pathway is activated and phosphorylates transcription factors (TFs), which can activate expression of pathogenesis-related (PR) genes as well as the fluorescent protein reporter. P, phosphate; Ter, terminator. (*Source*: Reproduced with permission of Liu et al. (2013b).)

A pioneering designer circuit was developed by Antunes et al. (2011) for the phytosensing of explosives, in which, applied 2,4,6-trinitrotoluene (the explosive TNT) was sensed by computationally designed artificial periplasmic TNT receptors in plants (Fig. 17.8a). Inputs from the interaction between TNT and receptors were linked to the histidine kinase (HK) signaling pathway; components of which are evolutionally conserved between bacteria and plants. Once the HK pathway was activated in engineered plants, the bacterial response regulator PhoB was translocated to the plant nucleus to activate plant gene expression. On TNT binding, the receptors bound an engineered transmembrane fusion protein Fls-Trg-PhoR, which then phosphorylated the bacterial response regulator PhoB-VP64. The translocation of phosphorylated PhoB-VP64 into the plant nucleus then activated an engineered de-greening gene circuit driven by a synthetic promoter PlantPho. Thus, upon ligand binding and receptor activation, the synthetic circuits caused rapid chlorophyll loss, allowing a visible (bleached) leaf color phenotype when TNT was present.

In a second example, a phytosensing system for bacterial pathogens was generated by Liu et al. (2011, 2013a). Here, a fluorescent reporter protein gene was driven by various forms of synthetic promoters containing plant-inducible regulatory elements to plant signal defense molecules (i.e., salicylic acid, ethylene, and jasmonic acid) (Fig. 17.8b). Following pathogen infection, the plant defense signaling pathways was activated, which, in turn, led to phosphorylation of MAP kinases and then corresponding transcription factors that bind to those inducible regulatory elements. Activation of expression of the reporter gene driven by each synthetic promoter demonstrated that these synthetic promoters were capable of responding in predictable ways to different bacterial pathogens and respective hormones in transgenic plants (Fig. 17.8b). These bacterial phytosensing plants

were demonstrated to be effective in field experiments, which is one of the first field-deployments of plant synthetic biology (Fethe et al. 2014).

These pioneering examples demonstrate the potential of phytosensors that will be useful in agriculture and other applications. These engineered phytosensors can be used as wide-area early-warning sentinels to indicate the presence of environmental toxic chemicals or plant pathogens.

17.5. CONCLUSIONS

The small but growing community of researchers in plant synthetic biology is applying the principles and methodologies of microbial synthetic biology into plant synthetic biology, even though this progress is currently slow and costly. Biological components are sometimes context-dependent and may not be easily predictable when ensembled into complex organisms such as plants. Efforts are needed to develop and characterize parts (phytobricks), and improve their deployment into plants. However, such efforts have raised other important challenges for plant synthetic biology, particularly assuring biosafety and the uncertainties that surround how synthetic biology will be regulated. Possibly, synthetic biology can address some biosafety concerns, but it might amplify existing concerns as increasingly larger amounts of DNA and protein are incorporated into crops.

LIFE BOX 17.1. JOSHUA YUAN

Joshua Yuan, Associate Professor, Systems Biology and Bioenergy, Texas A&M University.



Joshua Yuan by the microarray printer at the Gallo Center of UCSF. Courtesy of Joshua Yuan.

I became interested in plant research during my college years when I worked as an undergraduate assistant with Prof. Pifang Zhang in Fudan University, Shanghai, China. After graduation, I was enrolled as a master's student in the University of Arizona, where I met many elite plant biologists including David Galbraith, who later became the advisor for my master's thesis studying the expression of ice plant water channel promoters in Arabidopsis and developing the Microarray Analysis of Nuclear TRAnscriptome (MANTRA) technology.

David was a great mentor who always gives a grace period to allow students to grow as scientists. After the master's degree, I took an adventure into industry to work at the new BASF Plant Sciences LLC, where I helped the company to established functional genomics platforms. I quickly found that my nature of curiosity and desire for free-style research doesn't fit an industry career well, and moved back to academia to work at the University of California, San Francisco (UCSF) as a microarray manager, where I helped with different neuroscience projects. Regardless of being accepted by several top graduate programs, my attempt to go back to graduate school failed due to the complicated issue of permanent residency application.

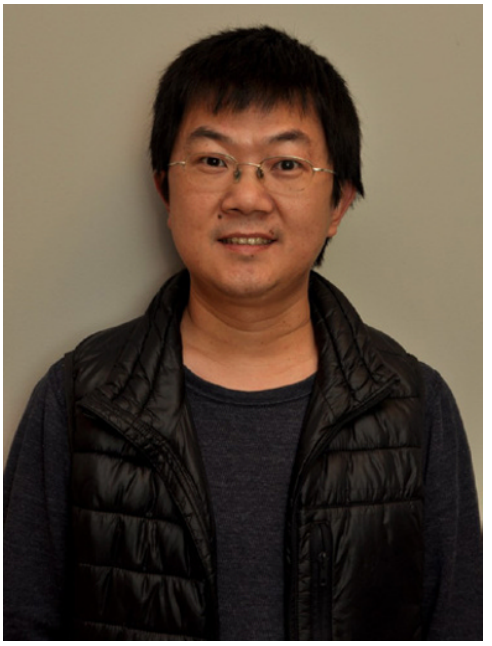
However, I was lucky enough to be offered a job at University of Tennessee (UT) to manage their genomics hub and pursue a PhD degree at the same time. I got a chance to work with people such as Neil Rhodes, Neal Stewart, and Feng Chen in my job there and

graduate studies. UT turned out to be a promised land for me, where I had a chance to revive my love for plant biology research. Neal is a great mentor, who always encourages you to go beyond your limits to develop a multidisciplinary research interest emphasizing on both fundamental research and application. My research at UT covered a broad spectrum, ranging from identifying volatile producing genes involved in tritrophic interaction, discovering the genes for low-temperature germinability, genetic engineering of key cell wall genes for better bioenergy feedstock, to developing bioinformatics tools for genomics data analysis. I was lucky enough to be trained by scientists with strong background[s] in technology development as well as both basic and applied scientific research, which makes me believe that my research should be driven by new technology, scientific questioning and needs of the society. After wandering in different fields, I came to realize that plant biotechnology is emerging as a field with more and more significant impact on our society and lives. As a traditional source of food, energy, and pharmaceuticals for mankind, the success in plant biotechnology research will enable more environmentally friendly energy supplies, more food, better nutrition, and cheaper healthcare products, all of which will contribute to the sustainable growth and peaceful development of human society.

After my time at UT, I became an assistant professor and now associate professor at Texas A&M University where I continue to explore basic research that has applications. My research spans systems biology and synthetic biology of plants, which, together, are interesting fields to combine that will greatly help the world.

LIFE BOX 17.2. WUSHENG LIU

Wusheng Liu, Research Assistant Professor, University of Tennessee.



Wusheng Liu. Courtesy of Wusheng Liu.

Born and raised in a small village surrounded with mountains, plants, and animals in Anhui Province, China, I was much influenced by agricultural environment and loved to take care of plants and animals. Moving from southern to northeastern China, I attended Northeast Forestry University in Harbin, China, for my B.A. degree in landscape architecture. About half of my classes were required to be plant science while the other half were in design. I felt more passionate about subjects in plant science, especially when I met my professors Drs. Lihuan Zhuo and Puhua Huang. Following a 2-month field trip with both of them to northeastern China, I decided to pursue my MS degree in botany under their supervision by studying plant morphological variations and population genetics. After graduation, I became an assistant lecturer, and then a lecture in plant science at Beijing Forestry University. At that time, my interest developed specifically in plant molecular biology. Benefitting from my collaboration with Dr. Shiliang Zhou at the

Institute of Botany, Chinese Academy of Sciences, I decided to apply to US universities for pursuing my doctoral degree. With some luck, I was accepted in Dr. Randall Small's lab at UT and studied molecular evolution of MADS-box genes in cotton for my PhD dissertation. Randy was a great mentor who encouraged me to take many classes and read tons of papers, and always gave me lots of grace time to grow as a scientist. All of these experiences together really built up a good base for future self-education.

After a short period of postdoc training in Dr. Elena Shpak's lab at UT, I went just across campus to join Dr. Neal Stewart's lab as a postdoc. Neal is really a great scientist and mentor, who always encouraged me to go beyond my upper limits to pursue answers to questions related to fundamental and applied science.

Here I became interested in plant synthetic biology, which currently is a growing and fascinating research area, which combines engineering principles with plant molecular biology. Its goal is to redesign a system for a particular purpose and to better understand biology by reconstruction. I just embraced this unique opportunity. My research projects combine plant molecular biology, genetics, bioinformatics, and synthetic biology, ranging from *de novo* soybean cyst nematode (SCN)-inducible motif discovery, bacterial pathogen phytosensing, targeted (trans)gene activation using transcription activator-like effectors (TALEs) from *Xanthomonas* and the binary Q expression system from *Neurospora*, to genome editing using clustered regularly interspaced short palindrome repeats (CRISPR). I have a strong desire to conduct research focusing on some economically important crops or bioenergy feedstocks. I especially enjoy the use of "omics," particularly transcriptomics and bioinformatics, and synthetic biology to pursue answers to some basic research questions in science. I hope what I discover in research will provide better understanding and more insights for improvements of crops/feedstocks qualities.

REFERENCES

Ainley WM, Sastry-Dent L, Welter ME, Murray MG, Zeitler B, Amora R, Corbin DR, Miles RR, Arnold NL, Strange TL, Simpson MA, Cao Z, Carroll C, Pawelczak KS, Blue R, West K, Rowland LM, Perkins D, Samuel P, Dewes CM, Shen L, Sriram S, Evans SL, Rebar EJ, Zhang L, Gregory PD, Urnov FD, Webb SR, Petolino JF (2013): Trait stacking via targeted genome editing. *Plant Biotechnol J* **11**:1126–1134.

Alterovitz G, Muso T, Ramoni MF (2009): The challenges of informatics in synthetic biology: from biomolecular networks to artificial organisms. *Brief Bioinform* **11**:80–95.

Anderson JC, Clarke EJ, Arkin AP, Voigt CA (2006): Environmentally controlled invasion of cancer cells by engineering bacteria. *J Mol Biol* **335**:619–627.

Andrianantoandro E, Basu S, Karig DK, Weiss R (2006): Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**:2006.0028.

Antunes MS, Morey KJ, Tewari-Singh N, Bowen TA, Smith JJ, Webb CT, Hellinga HW, Medford JI (2009): Engineering key components in a synthetic eukaryotic signal transduction pathway. *Mol Syst Biol* **5**:270.

Antunes MS, Morey KJ, Smith JJ, Albrecht KD, Bowen TA, Zdunek JK, Troupe JF, Cuneo MJ, Webb CT, Hellinga HW, Medford JI (2011): Programmable ligand detection system in plants through a synthetic signal transduction pathway. *PLoS One* **6**:e16292.

Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekraspv V (2015): Editing plant genomes with CRISPR/Cas9. *Curr Opin Biotechnol* **32**:76–84.

Benner SA (2003): Act natural. *Nature* **421**:118.

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005): Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**:2551–2561.

Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, Rock JM, Lee YL, Garrison R, Schulenberg L, Blue R, Worden A, Baker L, Faraji F, Zhang L, Holmes MC, Rebar EJ, Collingwood TN, Rubin-Wilson B, Gregory PD, Urnov FD, Petolino JF (2009): Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol Biol* **69**:699–709.

Chan LY, Kosuri S, Endy D (2005): Refactoring bacteriophage T7. *Mol Syst Biol* **1**:2005.0018.

Dymond JS, Richardson SM, Coombes CE, Babatz T, Muller H, Annaluru N, Blake WJ, Schwerzmann JW, Dai J, Lindstrom DL, Boeke AC, Gottschling DE, Chandrasegaran S, Bader JS, Boeke JD (2011): Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**:471–476.

Ellis T, Wang X, Collins JJ (2009): Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* **27**:465–471.

Elowitz MB, Leibler S (2000): A synthetic oscillatory network of transcriptional regulators. *Nature* **403**:335–338.

Endy D (2005): Foundations for engineering biology. *Nature* **438**:449–453.

Fethe MH, Liu W, Burris JN, Millwood RJ, Mazarei M, Rudis MR, Yeaman DG, Dubosquielle M, Stewart CN Jr (2014: The performance of pathogenic bacterial phytosensing transgenic tobacco in the field. *Plant Biotechnol J* **12**:755–764.

Gardner TS, Cantor CR, Collins JJ (2000): Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**:339–342.

Gibbs WW (2004): Synthetic life. *Sci Am* **290**:74–81.

Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang, R-Y, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi Z-Q, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA III, Smith HO, Venter JC (2010): Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**:52–56.

Havens KA, Guseman JM, Jang SS, Pierre-Jerome E, Bolten N, Klavins E, Nemhauser JL (2012): A synthetic approach reveals extensive tenability of auxin signaling. *Plant Physiol* **160**:135–142.

Hobom B (1980): Surgery of genes. At the doorstep of synthetic biology. *Medizin Klinik* **75**:14–21.

Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987): Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169**:5429–5433.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012): A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**:816–821.

Kobayashi H, Karn M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ (2004): Programmable cells: interfacing natural and engineered gene networks. *Proc Natl Acad Sci USA* **101**:8414–8419.

Koschmann J, Machens F, Becker M, Niemeyer J, Schulze J, Bulow L, Stahl DJ, Hehl R (2012): Intergration of bioinformatics and synthetic promoters leads to the discovery of novel elicitor-responsive cis-regulatory sequences in Arabidopsis. *Plant Physiol* **160**:178–191.

Levskaya A, Chevalier AA, Tabor JJ, Simpson ZB, Lavery LA, Levy M, Davidson EA, Scouras A, Ellington AD, Marcotte, EM, Voigt CA (2005): Synthetic biology: engineering *Escherichia coli* to see light. *Nature* **438**:441–442.

Levskaya A, Weiner OD, Lim WA, Voigt CA (2009): Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**:997–1001.

Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013): Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol* **31**:688–691.

Liu W, Stewart CN Jr (2015): Plant synthetic biology. *Trends Plant Sci* **20**:309–317.

Liu W, Mazarei M, Rudis MR, Fethe MH, Stewart CN Jr (2011): Rapid *in vivo* analysis of synthetic promoters for plant pathogen phytosensing. *BMC Biotechnol* **11**:108.

Liu W, Mazarei M, Rudis MR, Fethe MH, Peng Y, Millwood R, Shoene G, Burris JN, Stewart CN Jr (2013a): Bacterial pathogen phytosensing in transgenic tobacco and *Arabidopsis*. *Plant Biotechnol J* **11**:43-52.

Liu W, Yuan JS, Stewart CN Jr (2013b): Advanced genetic tools for plant biotechnology. *Nat Rev Genet* **14**:781–793.

Liu W, Mazarei M, Peng Y, Fethe MH, Rudis MR, Lin J, Millwood RJ, Arelli PR, Stewart CN Jr. (2014): Computational discovery of soybean promoter *cis*-regulatory elements for the construction of soybean cyst nematode inducible synthetic promoters. *Plant Biotechnol J* **12**: 1015–1026.

Lu TK, Collins JJ (2007): Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci USA* **104**:11197–11202.

Lu TK, Collins JJ (2009): Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proc Natl Acad Sci USA* **106**:4629–4634.

Mahfouz MM, Piatek A, Stewart CN Jr (2014): Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnol J* **12**: 1006–1014.

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E (2005): Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**:174–82.

Nekrasov V, Staskawicz B, Weigel D, Jones JDG, Kamoun S (2013): Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**:691–693.

Pourcel C, Salvignol G, Vergnaud G (2005): CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**:653–63.

Purnick PEM, Weiss R (2009): The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* **10**:410–422.

Rawls R (2000): 'Synthetic biology' makes its debut. *Chem Eng News* **78**:49–53.

Ro D-K, Paradise EM, Quellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006): Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**:940–943.

Rollie S, Mangold M, Sundmacher K (2011): Designing biological systems: systems engineering meets synthetic biology. *Chem Eng Sci* **69**:1–29.

Schwille P (2011): Bottom-up synthetic biology: engineering in a tinkerer's world. *Science* **333**:1252–1254.

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013): Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **31**:686–688.

Slusarczyk AL, Lin A, Weiss R (2012): Foundations for the design and implementation of synthetic genetic circuits. *Nat Rev Genet* **13**:406–420.

Smith HO, Hutchison CA 3rd, Pfannkoch C, Venter JC (2003): Generating a synthetic genome by whole genome assembly: *ΦX*174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci USA* **100**:15440–15445.

Szostak JW, Bartel DP, Luisi PL (2001): Synthesizing life. *Nature* **409**:387–390.

Tabor JJ, Salis HM, Simpson ZB, Chevalier AA, Levskaya A, Marcotte EM, Voigt CA, Ellington AD (2009): A synthetic genetic edge detection program. *Cell* **137**:1272–1281.

Terns MP, Terns RM (2011): CRISPR-based adaptive immune systems. *Curr Opin Microbiol* **14**:321–327.

Wang Y, Yau Y-Y, Perkins-Balding D, Thomson JG (2011): Recombinase technology: applications and possibilities. *Plant Cell Rep* **30**:267–285.

You L, Cox RS 3rd, Weiss R, Arnold, FH (2004): Programmed population control by cell-cell communication and regulated killing. *Nature* **428**: 868–871.

INDEX

ABA *see* abscisic acid
ABC model of flower development, 79, 93–94
abscisic acid, 78, 90–96, 110, 114, 238
Agrobacterium tumefaciens-mediated transformation
 analysis of transgenic plants, 285–286, 294–305
 history, 16–17
 intellectual property, 352–353
 methodologies, 216–251, 262–286
 vectors, 187–202
agroinfiltration, 271, 284, 393
anther culture, 119
antibiotics resistance gene *see* NptII and Hpt
anti-commons, 352–355
antisense construct, 243, 248
apomixis, 35–36, 71
Arnold, Michael, 40–42
asilomar conference on biosafety, 313
autogamy *see* self-fertilization
auxin, 82–99, 110–112, 115, 121–123, 179, 188, 240, 244, 246, 393–395

Bacillus thuringiensis see Bt transgenes
backcross breeding, 38, 61–62
Baenziger, Stephen, 74–75
bar gene *see pat* gene
barnase, 246–247
barstar, 246–247
Bartsch, Detlef, 344–346
Bayh-Dole Act, 351
Bennett, Alan, 360–361
beta glucuronidase *see* GUS reporter gene
binary vector *see Agrobacterium tumefaciens*, vectors
biofuels, 102, 205, 208, 226–227, 256–258
bioinformatics, 156–163, 167–169, 172–176, 214, 389, 393–396
biolistics *see* particle bombardment
biosafety, 311–326, 334, 342
Borlaug, Norman, 14–15, 38, 74, 288, 379, 380
brassinolide, brassinosteriod, 86, 94, 98, 110, 114
Bt transgenes, 3, 190–191, 203, 218–220, 319–320, 327, 334–342, 373–376
Buell, C. Robin, 177–178
Bulk breeding, 60–64

CAAT box, 140, 190–191, 236
callus, 110–123, 241, 267
CAMBIA, 353, 359
CaMV 35S promoter, 199–203, 216, 228, 235–236, 239–255, 354, 372
Cardoza, Vinitha, 128–129
cauliflower mosaic virus *see* CaMV 35S promoter
cDNA, 166–167, 172, 192, 195–196, 201, 213, 301–302, 351
cell suspension culture, 117–119, 271
central dogma, 133, 138, 155
centromere, 21–22, 28–30, 163
Chilton, Mary-Dell, 15–18, 267, 288, 290
chloroplast transformation, 204–205, 275
Chrispeels, Maarten, 150–151
chromatin, 83, 88, 134–137, 141–144
cis-acting element, 161, 173, 175, 187, 204, 234
cisgenic, 201–202, 316, 380, 389, 396
cleistogamy, cleistogamous flowers, 33
clustered regularly-interspaced short palindromic repeats *see* CRISPR
codons, 70, 144–149, 190–191, 198–202, 234
Collins, Glenn, 125–126
complementary DNA *see* cDNA
computer-aided design, 389
Conner, Tony, 378–379
co-transformation, 283
Cre-*lox* recombination system, 253, 391
CRISPR, 204, 304–305, 316, 390, 393–394
cytokinin, 82–83, 95–98, 108–121, 188, 240, 246, 251

Delmer, Deborah, 100–102
development of
 flowers, 92–93
 fruits, 83–84

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.

© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

leaves, 89–92
roots, 88–89
Dhillon, Taniya, 257–258
2,4-dichlorophenoxyacetic acid (2,4-D), 112
disease-resistance, 47, 51–52, 56–58, 61, 67–69, 93, 96–99, 220–222
Dixon, Richard A., 39–40
DNA blot analysis *see* Southern blot analysis
DNA synthesis and assembly, 384, 388
doubled haploids, 68–69, 119

EFSA, 316–317, 322, 327
electroporation, 182, 278, 282–283
ELISA, 303
embryogenesis (zygotic), 80–84
enhancers, transcriptional, 140–141, 200, 235, 243
enhancer trap, 249
5-enolpyruvylshikimate 3-phosphate synthase *see* EPSPS
environmental impact quotient, 8–10
Environmental Protection Agency, U.S. *see* EPA
environmental risk assessment (ERA), 334–339
enzyme-linked immunosorbent assay *see* ELISA
EPA, 313, 317–319, 324
epigenetics, 53, 117, 143
epistasis, 49
EPSPS, 215–216, 224, 244–246
ethylene, 95–99, 109–110, 120, 124, 396
European Food Safety Authority *see* EFSA
event, transgenic, 68, 71, 182, 204, 216, 242, 246, 272, 275, 283, 285, 294–305, 315, 338, 341, 373
exon, 139–144, 166–167, 172, 351, 388
explant, 108–124

FDA, 313–322, 375
Finer, John, 287–288
Flavr Savr tomato, 314, 368, 374–375
floral dip transformation, 271–272, 277, 283–284, 286, 294, 317
FLP-*FRT* recombination system, 253, 390–391
fluorescence *in situ* hybridization (FISH), 31
Food and Drug Administration, US *see* FDA
Fraley, Robert T., 17–18, 290, 314
Frankenstein, 366–367
Freedom to operate (FTO), 355–359

GA *see* gibberellic acid
gametogenesis, 80–82
Gateway cloning, 192–207
gene gun *see* particle bombardment
gene silencing *see* antisense construct; RNA interference (RNAi)
genetic linkage, 26, 49, 61, 69, 162, 251
genome, 21–38, 69–73, 142, 156, 162, 172–176, 192–197
genome editing, 389–393
genomics, 162–166, 212–213, 243, 256
genotype, 23–26, 33–38, 43–69
GFP reporter gene, 198, 202, 241, 250–253, 265, 272, 276–277, 285, 300–303
gibberellic acid, 85, 95–98, 110, 114, 203
Gibson Assembly, 192, 195–197
Gibson Hummingbird, 301
glyphosate resistance, 8–13, 215–217, 241, 244–245 *see also* herbicide resistance
Golden Rice, 222–224, 227, 229–231, 270, 315
Gonsalves, Dennis, 221–222, 227–229
green fluorescent protein *see* GFP reporter gene
GUS reporter gene, 89, 241, 246–249, 285

Hemizygous, 300–301
Herbicide resistance (or tolerance), 2–13, 214–220, 241–245, 265, 284, 315, 317–319, 376
heterologous, 118, 235, 354, 356
Hinchee, Maud, 361–362
histone acetyltransferase, 141–143
histones, 134, 139, 141–143
horizontal gene transfer, 254, 353
HPT selectable marker gene, 241–244, 353
hybrids, 25–27, 31–38, 44, 48, 50, 53, 56–58, 65
hybrid varieties, 57–58, 65–67
hygromycin phophotransferase *see* HPT selectable marker gene

IAA, 89, 97, 108–112, 123, 240, 393–395
IBA, 110, 112
indole acetic acid *see* IAA
indole butyric acid *see* IBA
insect resistance, 3–13, 218–220
intellectual property *see* patents
intron, 139–144, 163, 167, 172, 190, 200
IPT gene, 242, 246, 251
isopentyl transferase gene *see* IPT gene

jasmonic acid, jasmonate, 95–99, 110, 114, 396

kanamycin resistance gene *see* NptII selectable marker gene
Khush, Gurdev, 72–73
Klein, Ted, 273, 286–287
Kozak sequence, 146, 190

laser micropuncture, 279
left border, T-DNA, 188, 269, 289–290
ligase, 182, 184, 186, 192, 196, 197
Liu, Wusheng, 398–399
luciferase reporter gene, 153, 241, 247–250

male sterility, 34–35, 56, 246
Mann, David, 207–208

marker-assisted selection, 68–69, 214
marker-free strategies, 234, 250–255
marker gene, selectable, 189–190, 194, 198, 202, 205, 239–254, 284, 300, 305, 352–354
mass spectroscopy, 162, 168–176, 214
material transfer agreement (MTA), 350
matrix attachment region, 204
McHughen, Alan, 328–329
meiosis, 23, 27–32, 35–37, 57, 68, 80–81
Mendel, Gregor, 20–21, 23–27
Mendelian
 segregation, 23–26, 242, 285, 294, 300–301
 trait, 23–26, 45
meristems
 root apical, 79, 88–89
 shoot apical, 79, 86–88
messenger RNA *see* mRNA
metabolomics, 155–157, 160, 165, 170–172, 176, 214
methylation, DNA, 142–144, 182, 316
microarray analysis, 166–167, 172–175, 213, 395
microRNA (miRNA), 143, 199
mitosis, 23–29, 31, 81, 134
monarch butterfly, 336, 339–340, 373
Moon, Hong S., 305–306
mRNA
 analysis, 301–302 *see also* northern blot analysis and real-time PCR
 production during transcription, 135–144
 stability, 143, 190, 214
MS medium, 108–110
multiple cloning site, 192–198
Murashige and Skoog medium *see* MS medium
mutation breeding, 68, 70–71, 352

nanofiber arrays, 208, 279–280
National Institutes of Health guidelines *see* NIH guidelines
neomycin phosphotransferase *see* NptII selectable marker gene
NIH guidelines, 313
northern blot analysis, 301–305
NptII selectable marker gene, 190, 241–244, 247, 249, 353
nucleosome, 134–135, 137, 141

OECD, 313, 324
oils, modified, 224–225
open access, 175, 263, 358–359
open reading frame, 197–200, 205, 216, 388
open source, 359
ORF *see* open reading frame
Organization for Economic Cooperation and Development *see* OECD
organogenesis, 115, 120–123, 246
origins of replication, 186–190
outcrossing, 33–34, 36–38, 58, 63–67, 204
Ow, David, 152–153

papaya ringspot virus, 221–222, 227–229
Parrott, Wayne, 206–207
particle bombardment, 17, 191, 216, 220, 224, 234, 262–263, 272–276, 279, 287, 294, 299
patents, 216, 230, 277, 281, 348–351
pat gene, 217, 241–245, 265
pathogen resistance *see* disease-resistance
PCR
 defined, 192–193, 295
 real-time, 295–302
 use in characterizing transgenic plants, 285–286, 295–296
 use in cloning and vector construction, 194–196
pedigree breeding, 58–59
PEG, 120, 279
pharmaceuticals, produced in genetically engineered plants, 119, 201, 225–226, 314, 318, 356, 359, 388
phosphinothricin *see* PPT
phosphinothricin acetyltransferase gene *see pat* gene
phosphoisomerase gene *see* PMI gene
photomorphogenesis, 85–86
PIPRA, 354, 358–359
plant-produced pharmaceuticals *see* pharmaceuticals, produced in genetically engineered plants
plasmids, 181–205, 234, 251, 265, 268–269, 276, 299
plastid vector *see* chloroplast transformation
PMI gene, 245
polyadenylation *see* polyA tail
polyA tail, 144, 167, 190
polyethylene glycol *see* PEG
polygenic traits, 25, 45–46, 56
polylinker *see* Multiple cloning site
polymerase chain reaction *see* PCR
polypeptide, 134–135, 144–150, 197
polyploids, 36–38, 50, 56
population genetics, 50
position effects, 248, 275
posttranscriptional gene silencing *see* RNA interference (RNAi)
posttranslational modification, 150, 168, 214
Potrykus, Ingo, 224, 229–231, 270, 288, 330, 381
PPT, 244, 330
Prakash, Channapatna, 379–381
prior art, 350, 356
promoter
 constitutive, 86, 189–190, 199, 202–203, 234–236, 239–245, 248, 249, 254–255, 304, 354, 384, 389
 inducible, 199, 203, 235–240, 246, 253, 304, 383, 389–396

synthetic, 235, 237, 239–240, 388–389, 394–396
tissue-specific, 140, 203, 235–237, 246, 304, 354, 389, 390
trapping, 247–249
protein blot analysis *see* western blot analysis
proteomics, 156–157, 160–161, 165, 168–170, 172, 174–176, 214
protoplasts, 118–120, 227–279, 282, 330, 393
Public Intellectual Property Resource for Agriculture *see* PIPRA
Punnett Square, 25–27
putative transformants, 283, 285, 294, 295, 299

quantitative PCR *see* real-time PCR
quantitative trait loci (QTL), 50, 59, 68–70, 76

Raikhel, Natasha, 102–103
reading frame, 144–147, 166, 197–199
real-time PCR *see* PCR, real-time
recombination
homologous, 204–205, 275, 390
site-specific, 153, 182, 192–194, 203–205, 253, 383, 390–391 *see also* Gateway cloning and Cre-*lox*
recurrent selection, 53, 54, 56, 61, 64–66
regulatory flaws, 323–327
Reichert, Nancy A., 308–309
reporter genes, 89, 192, 198, 200, 239–254, 285, 293–295, 300, 305, 384–386
restriction endonuclease, 182–198, 285, 296–300
restriction enzyme *see* restriction endonuclease
ribonuclease, 247
right border, T-DNA, 188, 269, 289–290
RNA blot analysis *see* northern blot analysis
RNA interference (RNAi), 143, 199, 204, 275, 294, 305, 316
RNA polymerase II, 136–141, 190–191, 202, 235
RNA-Seq, 162, 166–168, 213–214
RNases, 143
Roundup, Roundup-Ready *see* glyphosate

salicyclic acid, 110, 225, 240, 396
seed germination, 80–83, 85–87, 94, 96, 123
segregation analysis, 300–301
selectable marker genes, 189–190, 194, 198, 202, 205, 233, 239–246, 249–255, 284, 294, 300, 305, 353–355
self-fertilization, 23–25, 33, 36, 48–50, 58, 63, 271
self-incompatibility, 32–35, 38, 81
self-pollination *see* Self-fertilization
Shelton, Tony, 343–344
shifting balance theory, 53, 55
Shillito, Ray, 329–331
signal transduction, 85–87, 90–91, 94–99, 161, 242, 387–388, 395
silicon carbide whiskers, 278
single-seed descent, 59–60, 69
somaclonal variation, 115, 117, 121, 123, 325, 372
somatic embryos, 113, 123, 204
Southern blot analysis, 166, 285, 296–303, 305
splicesome, 143–144
35S promoter *see* CaMV 35S promoter
stacked genes, 2, 56, 203–204, 218, 220, 236, 392
start codon, 146–148, 190
stop codon, 146–149, 198–199
surface sterilization of explants, 112
synthetic varieties, 64–67

TALE, TALEN, 204, 390, 392
Tanksley, Steven D., 75–76
tapetum, 80, 81, 246, 247
targeted transgene insertion, 204–206
TATA box, 190–191, 234–235, 239, 396
T-DNA, 187–189, 191, 202–204, 234–235, 244, 246, 251, 265–272, 281–283, 289–290, 299, 300, 353
Ti plasmid, 16–17, 188, 190, 240, 244, 246, 265, 268–269, 353
totipotency, totipotent, 108, 264
traits
input, 56, 214–222, 315
output, 222–227
transcription
factors, 82, 86–88, 91, 92, 96–99, 131, 140–141, 237–238, 254, 389, 396
process, 135–144
start site, 138–139, 190–191, 234–235
transcriptional activator-like effector *see* TALE, TALEN
transcriptomics, 156, 157, 166–167, 172, 388
transfer DNA *see* T-DNA
transfer RNA *see* tRNA
translation, 144–147
translation start site *see* start codon
tRNA, 147–149, 166
tumor-inducing plasmid *see* Ti plasmid

uidA gene *see* GUS reporter gene
United States Department of Agriculture
Animal and Plant Health Inspection Service
Biotechnology Regulation Service *see* USDA APHIS BRS
untranslated region *see* UTR
upregulated gene, 213
USDA APHIS BRS, 315, 318, 322, 324
UTR, 138, 190, 235

vaccines, produced in genetically engineered plants, 201, 225, 356, 359, 388
Van Montagu, Marc, 267, 289, 314
viral vectors, 278–279

virulence *(vir)* genes, 188, 234, 267–271, 278
virus-induced gene silencing (VIGS), 199

Wang, Kan, 289–290
western blot analysis, 302–303
Winkel, Brenda S.J., 103–105
World Trade Organization *see* WTO
Wright, Martha, 127–128
WTO, 321

Yang, Zhenbiao, 178–179
Yuan, Joshua, 387–388

Zinc finger nuclease (ZFN), 204, 316, 390–392

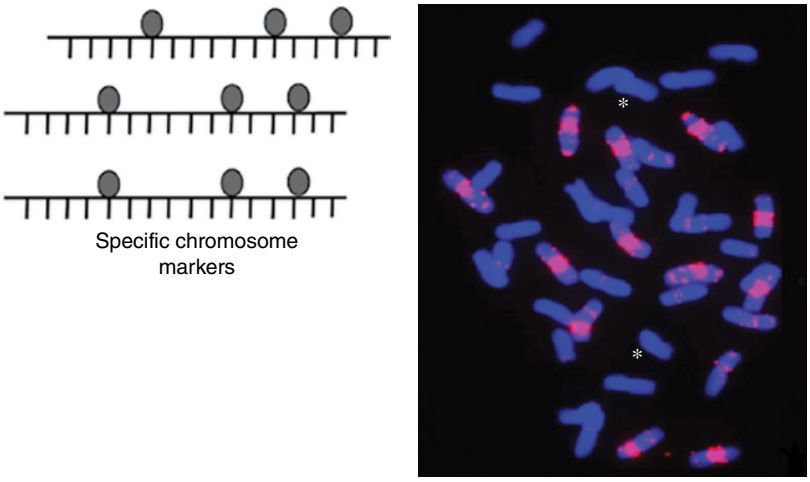


Figure 2.9. Fluorescence *in situ* hybridization (FISH) shows the physical location of a specific transgene or DNA.



Figure 3.17. Visualization of SNP markers on chromosome-1 for a set of soybean varieties. Each column represents a locus position on the chromosome, and each row represents a different soybean variety. Most loci have two alternate alleles, which are colored to represent the DNA base present in a homozygous state in the corresponding soybean variety. The predicted value of each allele is determined by testing a reference population where phenotypes are known. A predicted genotypic value of each soybean variety is then derived as a summation of predicted allele values, and varieties with the highest overall genotypic values are selected.

Plant Biotechnology and Genetics: Principles, Techniques, and Applications, Second Edition. Edited by C. Neal Stewart, Jr.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.

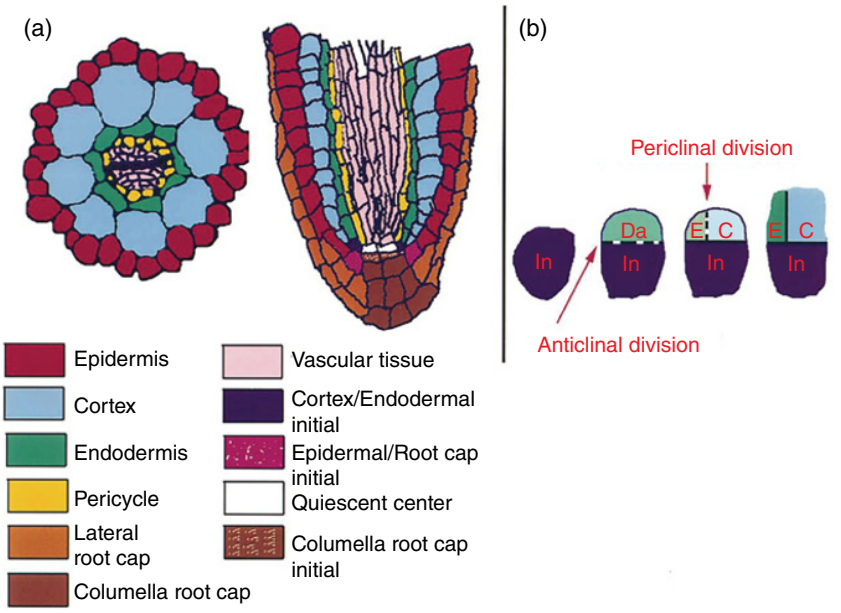


Figure 4.6. Root development. Arrangement (a) and division plane (b) of cell types within the developing root. (From Di Laurenzio et al. (1996). Reproduced with permission from Cell Press.)

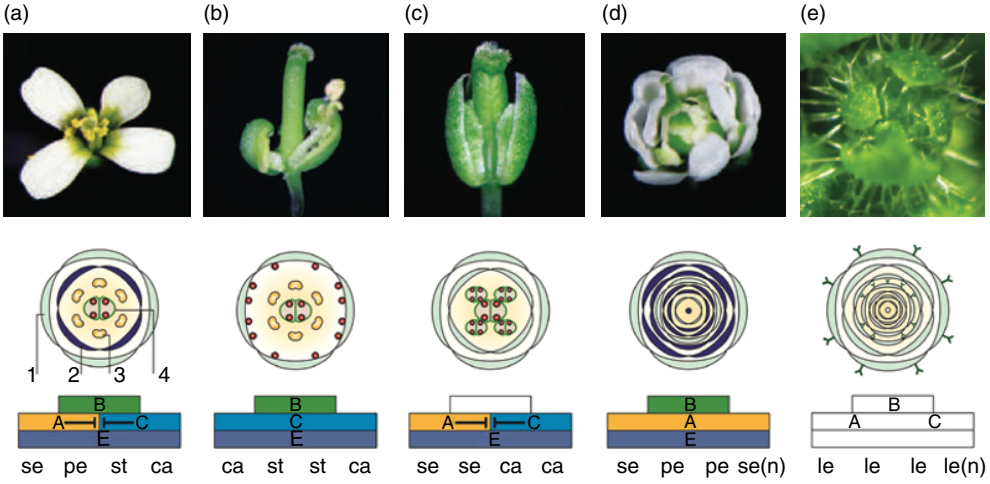


Figure 4.8. Flower development. *Arabidopsis* (a) wild-type, (b) *ap*2, (c) *pi*, (d) *ag*, and (e) *sep* flowers. Below each photo is a rendering of the ABC model as it functions in that flower. (From Krizek and Fletcher (2005). Reproduced with permission from Nature Publishing.)

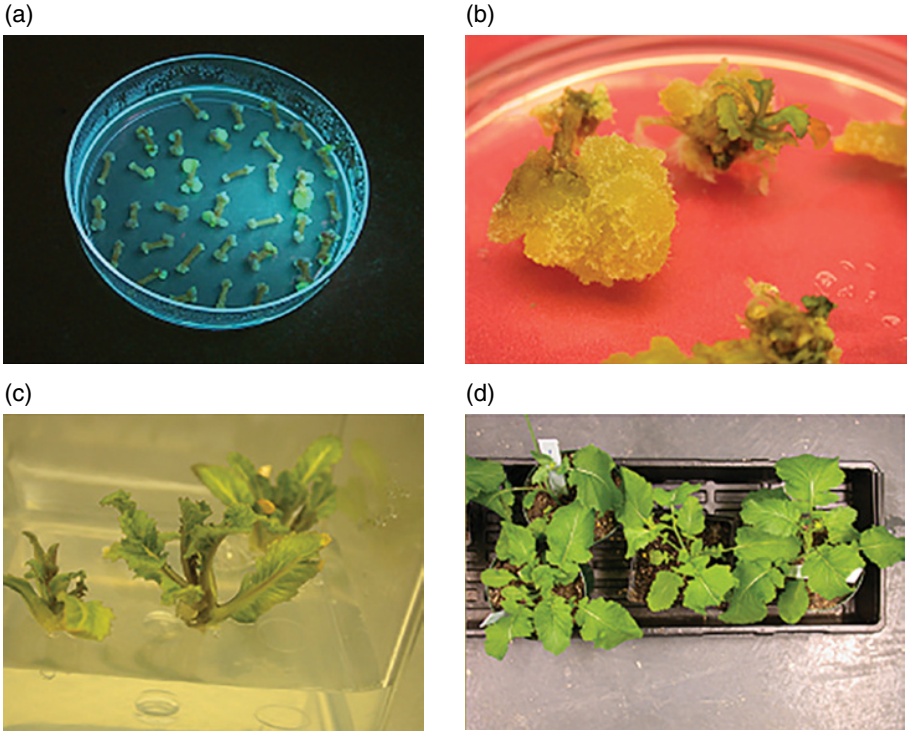


Figure 5.2. *Brassica juncea* plants produced from hypocotyls explants. Shoots are produced when a combination of auxin and cytokinin is used, which is a critical step. The key tissue culture stages for this system is (a) callus from hypocotyl explants; (b) shoots from callus; (c) elongating shoots; and (d) whole plantlets that have been transferred to pots.

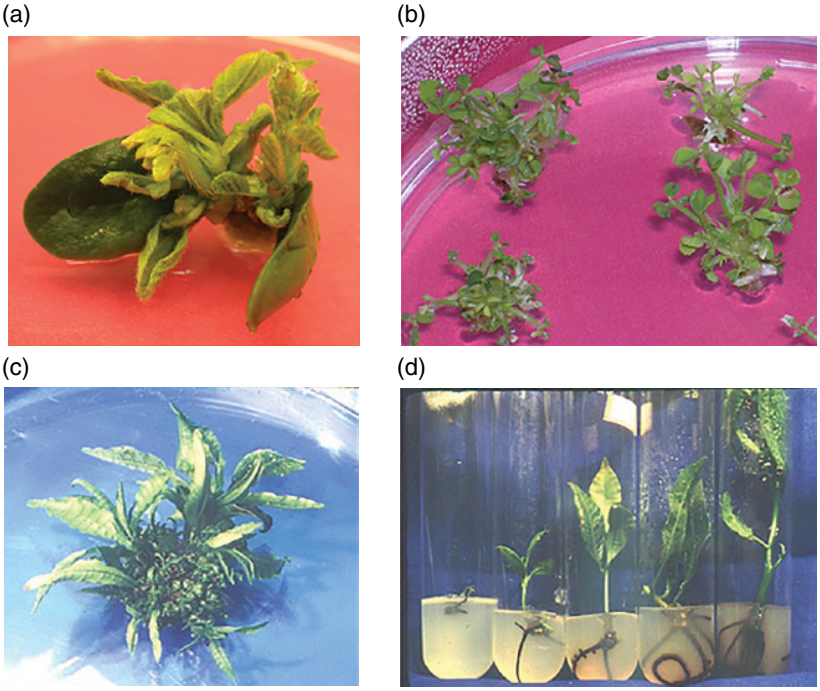


Figure 5.16. Several examples of direct organogenesis in various plant species: (a) multiple bud initiation from cotyledonary nodes of soybean, (b) shoot formation from multiple buds in *Medicago truncatula*, which is a relative of alfalfa, (c) shoot formation from multiple buds of cashew, and (d) the developments of roots and elongating shoots in cashew.

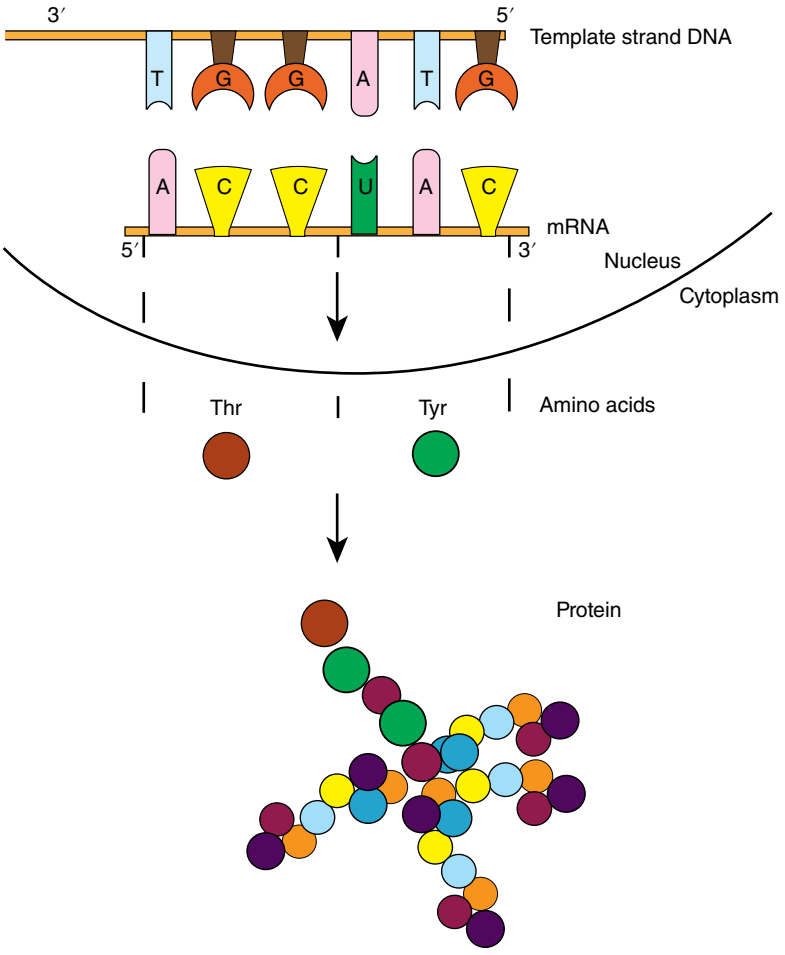


Figure 6.4. The central dogma: DNA is transcribed to RNA in the cell nucleus. RNA is translated to protein in the cell cytoplasm.

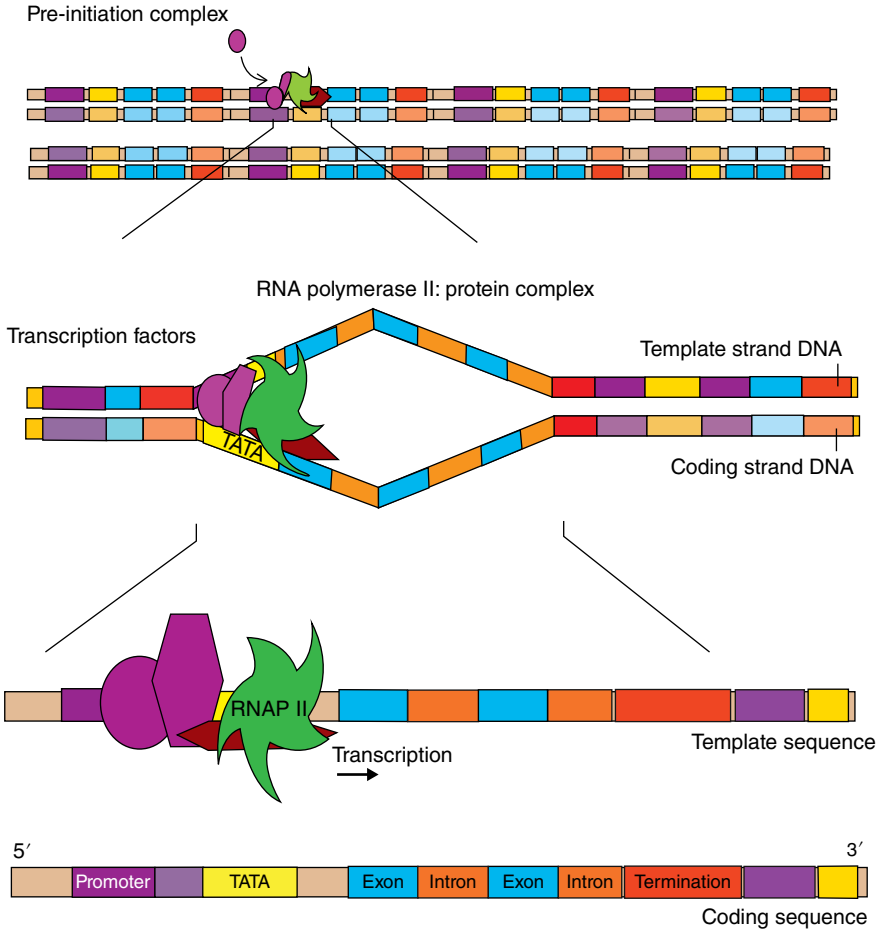


Figure 6.6. Overview of the early steps of transcription. A preinitiation complex is formed by a complex of transcription factors and RNA polymerase II (RNAP II). Association of the preinitiation complex with the start sequence (TATA) of the coding strand of DNA causes a conformation change and hydrogen bond breakage. This causes the DNA strands to separate so that transcription can proceed.

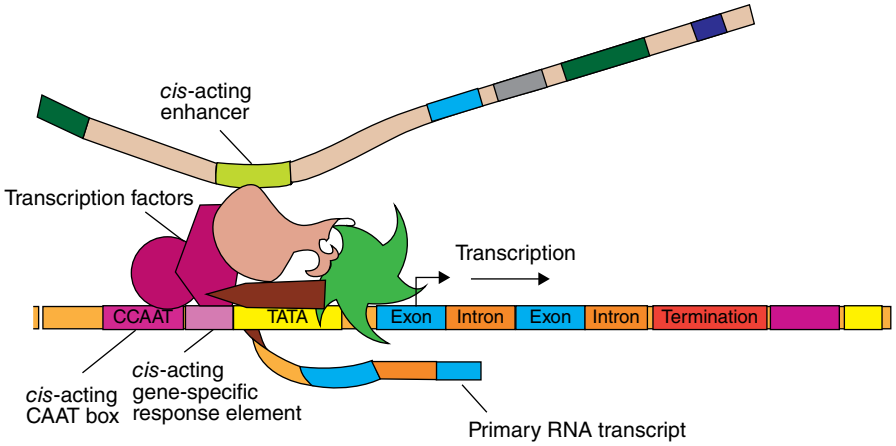


Figure 6.7. Regulation of transcription. The *cis*-acting elements are segments of DNA that regulate transcription; these segments may be adjacent to the gene such as the promoter (CAAT box) and the *cis*-acting gene-specific response elements, or they may be distant to the gene such as enhancers. The *trans*-acting elements are transcription factors and other regulatory proteins that may associate with the promoter, other proteins, or both.

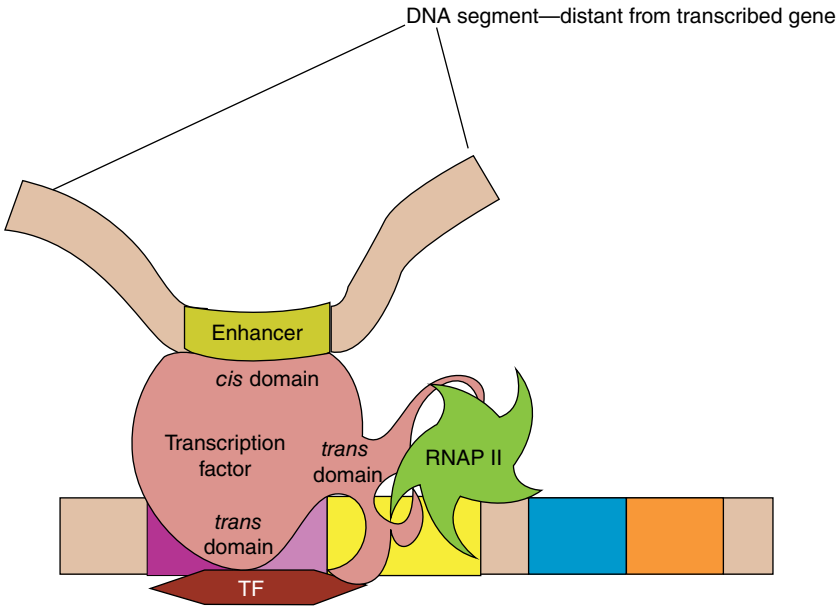


Figure 6.8. Transcription factors structure and function. Transcription factors may have domains that bind *cis*-acting elements such as enhancers, and domains that also bind *trans*-acting elements such as RNA polymerase (RNAP II) and other transcription factors.

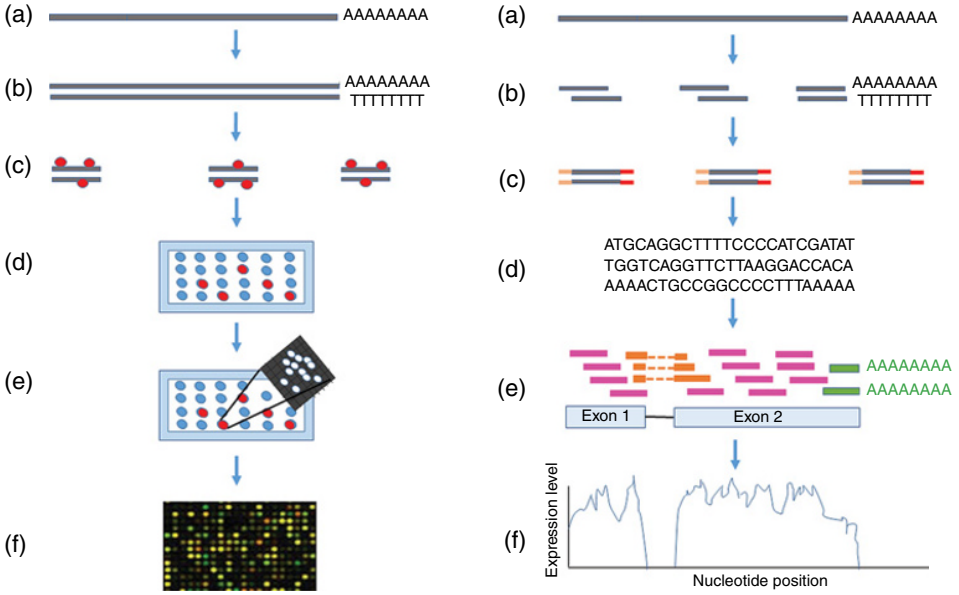


Figure 7.11. Comparison of the flow charts of microarray analysis (left) and RNA-seq (right). To conduct a microarray experiment, the following steps (shown on the left) are taken: (a) total RNA is extracted, (b) which is used for the template for cDNA synthesis, (c) followed by labeling and fragmentation, (d) hybridization and washing, (e) laser scanning, and (f) computer analysis of the expression profiles. RNA-seq shares steps or has analogous steps to microarray analysis, and shown to the right: total RNA is extracted (a) and is fragmented before or after cDNA synthesis (b), followed by ligation to adaptors (c), next-generation sequencing to produce huge amounts of short reads (d). These reads are mapped to a reference genome or transcriptome, or used for de novo assembly, and can be classified as junction reads, exon reads and poly(A) tail reads (e). Then, these reads are used to generate base-resolution expression profiles for different genes (f).

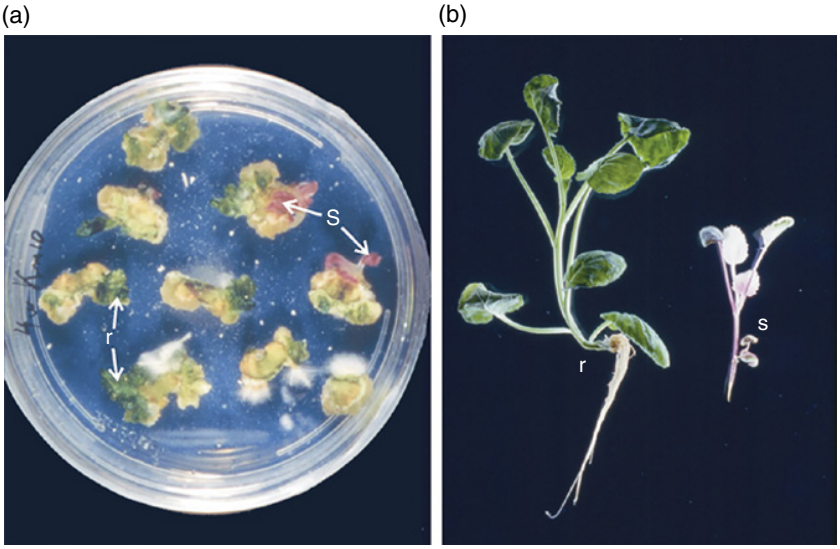


Figure 10.2. Selection of transgenic canola (*Brassica napus* cv Westar) on kanamycin-containing tissue culture media. Stem explants were first infected with an *A. tumefaciens* strain harboring a transformation vector with a chimeric *nptII* gene designed to confer kanamycin resistance on transformed plant tissue. (a) After cocultivation of plant tissue with *Agrobacterium* allowing transformation to occur, the plant tissues were transferred to tissue culture media containing kanamycin for growth of callus tissue and shoot differentiation. Much of the non-transformed tissues turned white (see arrows pointing to "s") and stopped growing because they were sensitive to the antibiotic. Transformed tissues remained green and continued to grow and differentiate because they were resistant to kanamycin (see the arrows pointing to "r"). (b) Transgenic shoots that differentiated in the presence of kanamycin were excised from the callus and transferred to media for the regeneration of roots. Escapes that were not truly kanamycin-resistant were unable to regenerate roots in the presence of the antibiotic. (*Source*: Courtesy of Pierre Charest).

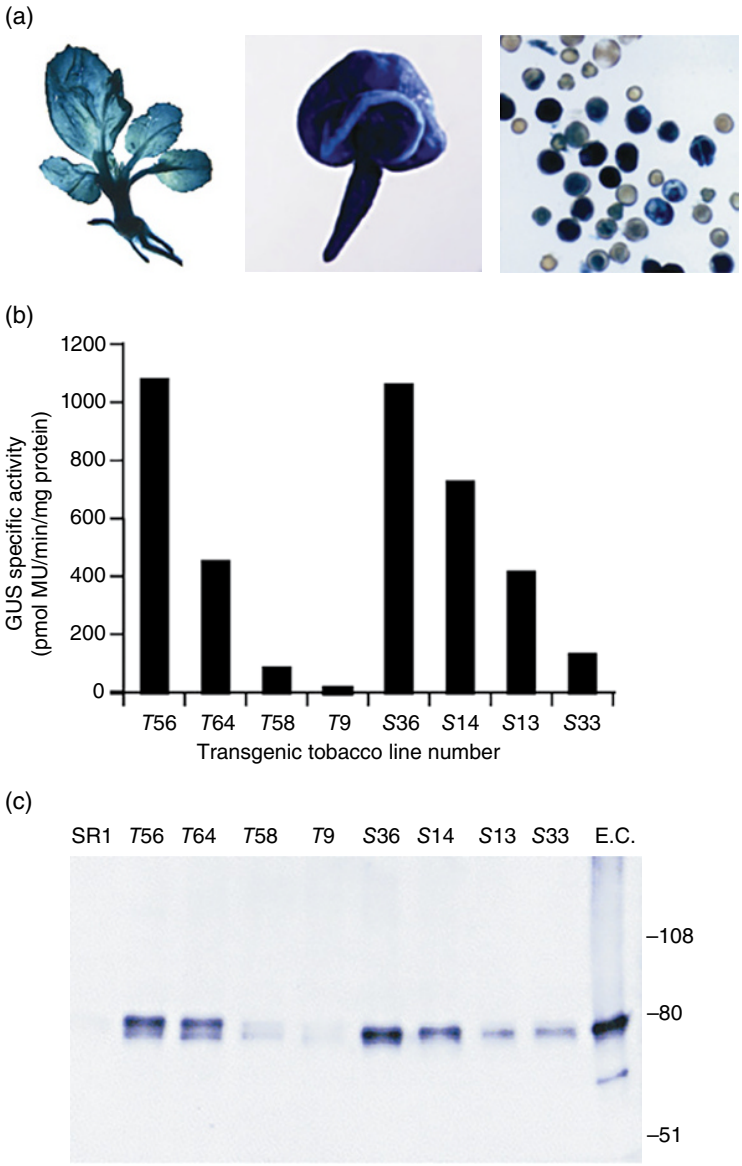


Figure 10.6. The *uidA* gene, coding for GUS, as an example of a reporter gene that has been extensively used in plants. (a) Histochemical staining for GUS activity using the substrate 4-methyl umbelliferyl glucuronide (MUG) allows detection of gene activity in specific tissues of transgenic plants. Shown in the figure are the staining of cauliflower plantlets in which constitutive expression of GUS is conferred by a strong constitutive promoter, tCUP; excised embryos from transgenic canola seeds in which seed-specific expression is conferred by the napin promoter; and transgenic canola pollen in which cell-specific expression is conferred by the pollen-specific (Bnm1) promoter. Note here that pollen cells are segregating as transformed and non-transformed cells indicated by the presence and absence of staining. (*Source*: Courtesy of Dan Brown.) (b) Measurement of GUS enzyme-specific activity using the substrate 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc). Each separate transgenic line of tobacco differs in the level of gene expression because of the variation in the influences on the inserted genes from the genetic elements and chromatin environment at the different sites of insertion. These are often called *position effects*. To compare differences among genes and elements introduced into transgenic plants, analyses must account for a large number of transgenic lines to reduce the influence of position effects. Reporter genes provide a valuable means for gathering large amounts of data. Here, a comparison of the promoter strengths of the 35S (plant lines with the *S* designation) and tCUP (plant lines with the *T* designation) constitutive promoters is inferred by comparing the activities of the reporter gene. (c) To ensure that the reporter gene reflects transcriptional activity, RNase protection assays are used to measure the relative amounts of GUS mRNA accumulating in the transgenic lines. This assay involves the formation of stable RNA duplexes with a radiolabeled antisense RNA probe followed by RNase digestion of the single-stranded RNA molecules so that the protected double-stranded RNA can be separated by gel electrophoresis and quantified.

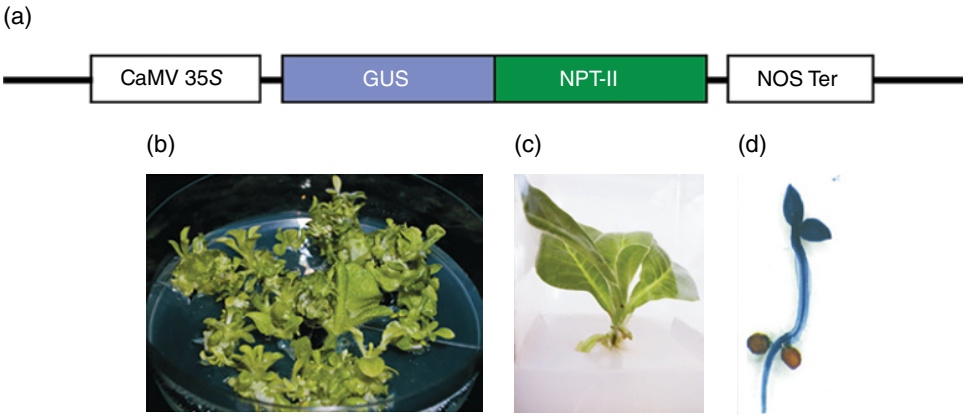


Figure 10.8. Fusion of a reporter and selectable marker gene to create a bifunctional gene: (a) GUS:NPTII fusion reporter system for plants that incorporates the *nptII* gene for kanamycin selection and the GUS reporter gene in a single module; (b) transformed tobacco shoots selected on kanamycin; (c) shoots with roots regenerated on kanamycin; and (d) a transgenic seedling after two generations showing retention of GUS gene activity indicated by the histochemical staining with the GUS substrate X-Gluc. (*Source*: From courtesy of Raju Datla.)



Figure 10.9. Luminescence detected in transgenic tobacco transformed with the firefly luciferase gene driven by the 35S promoter and watered with a solution of luciferin, the luciferase substrate. (*Source*: From Ow et al. (1986). Reproduced with permission of AAAS.)

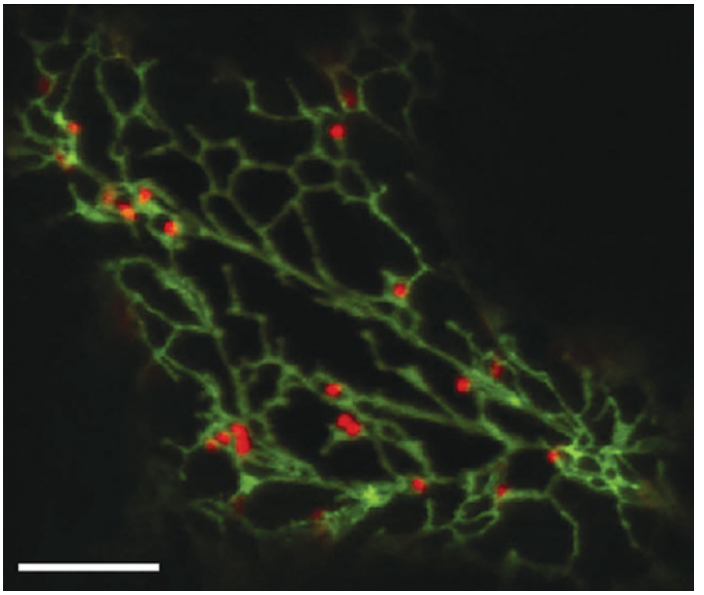


Figure 10.10. Confocal laser scanning microscopy of leaf mesophyll cells transiently expressing peptides fused to green fluorescent protein or GFP (green image) and yellow fluorescent protein (red image). GFP is fused to the HDEL tetrapeptide (spGFP-HDEL) to achieve ER retention and thus reveals the cortical ER network in leaf cells. The proximity of the Golgi to the ER network is revealed by the yellow FP fused to a Golgi glycosylation enzyme (ST-YFP). (Bar = 10 μm.) (*Source*: From Brandizzi et al. 2004.)

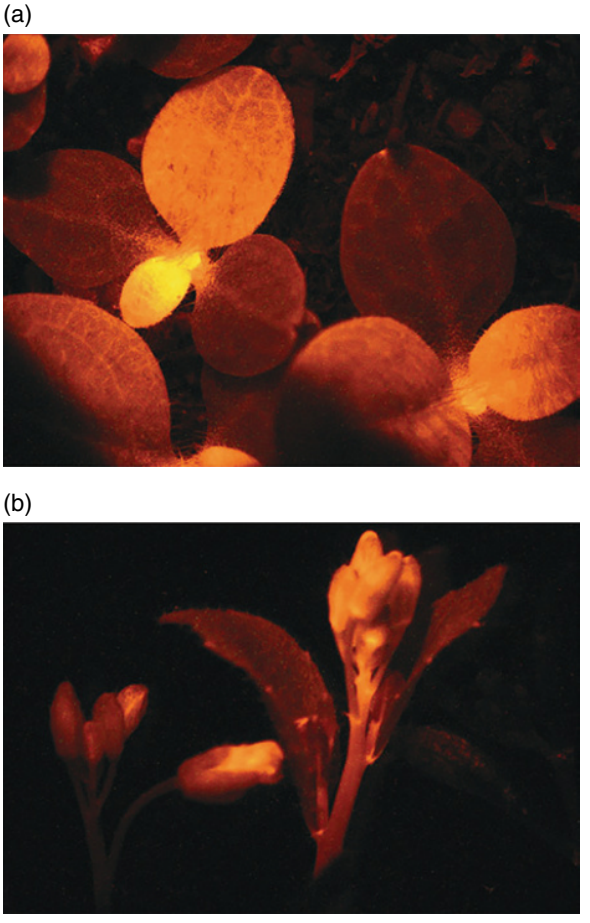


Figure 10.11. Orange fluorescent proteins whose genes were cloned from corals and expressed in tobacco (a) and *Arabidopsis* (b) plants.

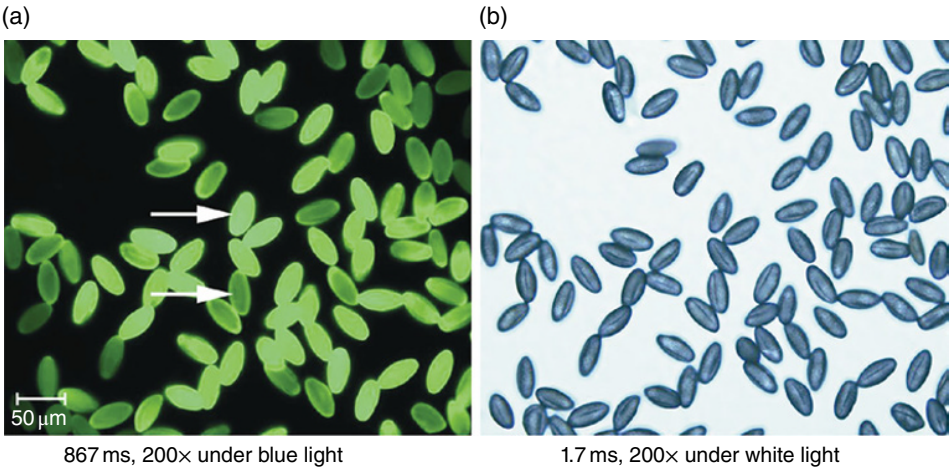


Figure 10.12. The green fluorescent protein (GFP) has been useful for marking whole plants using a 35S-GFP construct and plant parts such as pollen using GFP under the control of a pollen-specific promoter (Lat59) from tomato: (a) 867 ms, 200× under blue light and (b) 1.7 ms, 200× under white light. The arrows in (a) show GFP fluorescence of pollen cells. (*Source*: Courtesy of Moon & Stewart.)



Figure 11.5. Agroinfiltrated *Nicotiana benthamiana* plants showing high levels of GFP expression. The aerial parts of the tobacco plant were submerged in an *Agrobacterium* suspension and the plant was then placed under vacuum for infiltration. Courtesy of John Lindbo.

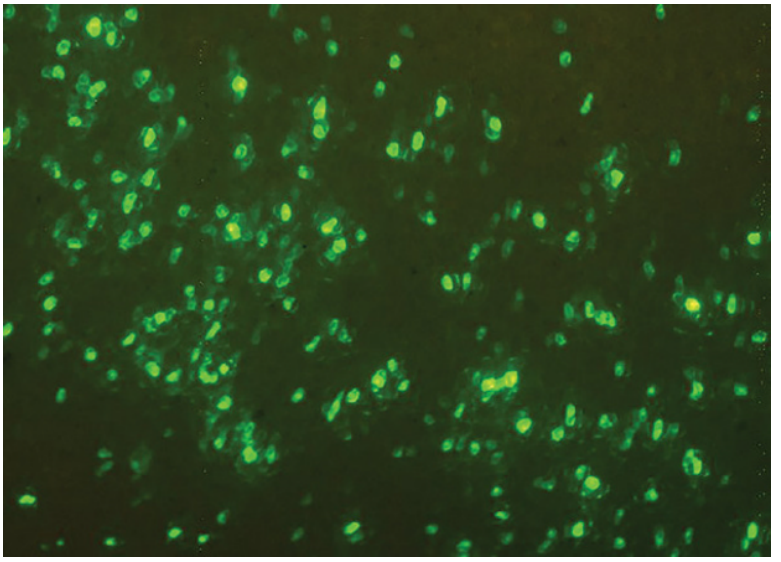


Figure 11.8. Particle bombardment-mediated transient GFP expression in lima bean cotyledonary tissues. This target tissue is flat, non-pigmented, and ideally suited for tracking GFP expression in individual transiently transformed cells.

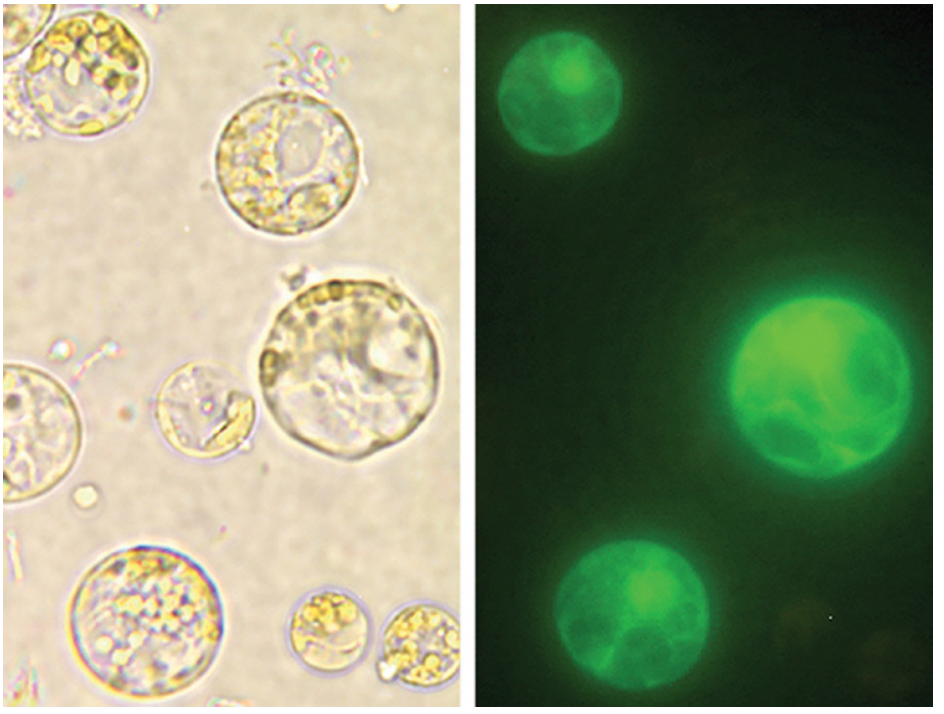


Figure 11.9. Maize protoplasts, electroporated with a *gfp* gene, showing bright field (left) and with GFP filters (right). Courtesy of JC Jang.

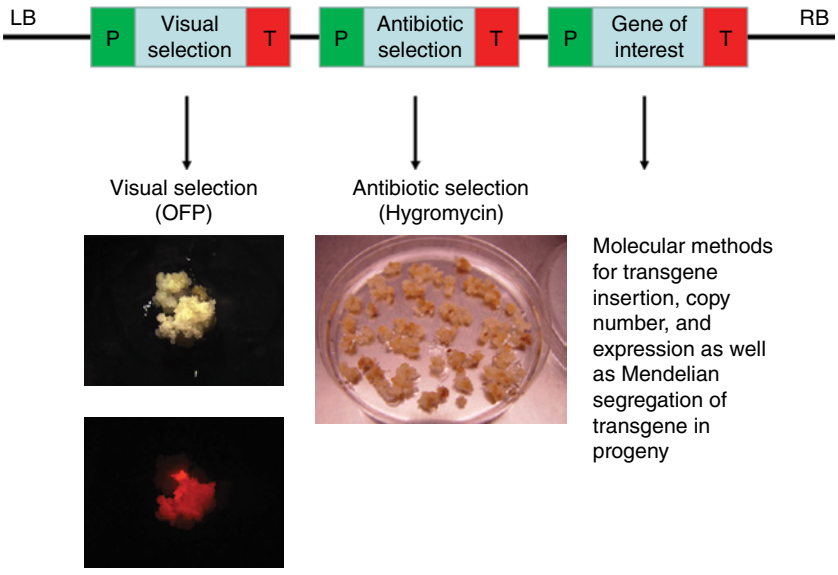


Figure 12.1. Overview of transgenic plant analysis. Several lines of evidence can be used together to assess whether the plants are truly transgenic and that the transgene of interest is expressed. Thanks to Mat Halter for assistance on this figure.

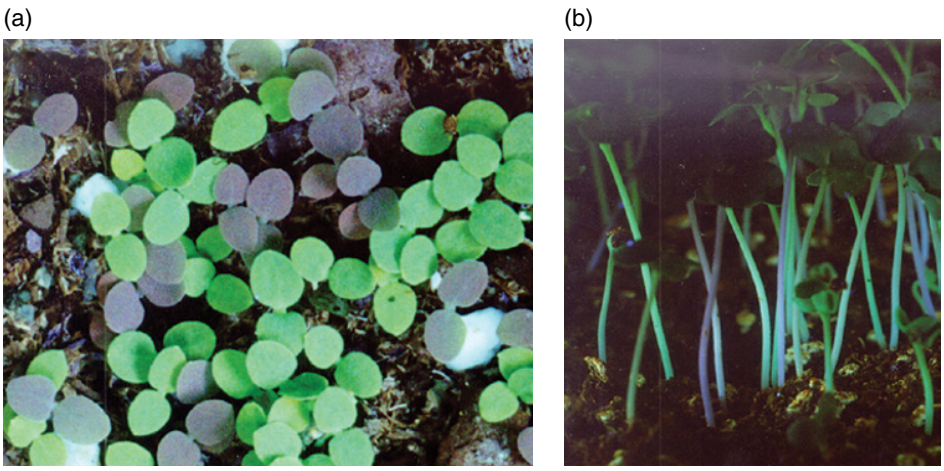


Figure 12.8. Segregation analysis of T1 transgenic (a) tobacco and (b) canola seedlings that have a single insert of a green fluorescent protein (GFP) gene. Under a UV light, the transgenic plants fluoresce green and the non-transgenic plants fluoresce red. The transgene presence and the single insert into the genome are confirmed by the Mendelian 3 : 1 segregation pattern in both of these cases. (*Source*: Reproduced with permission from Harper et al. (1999).)

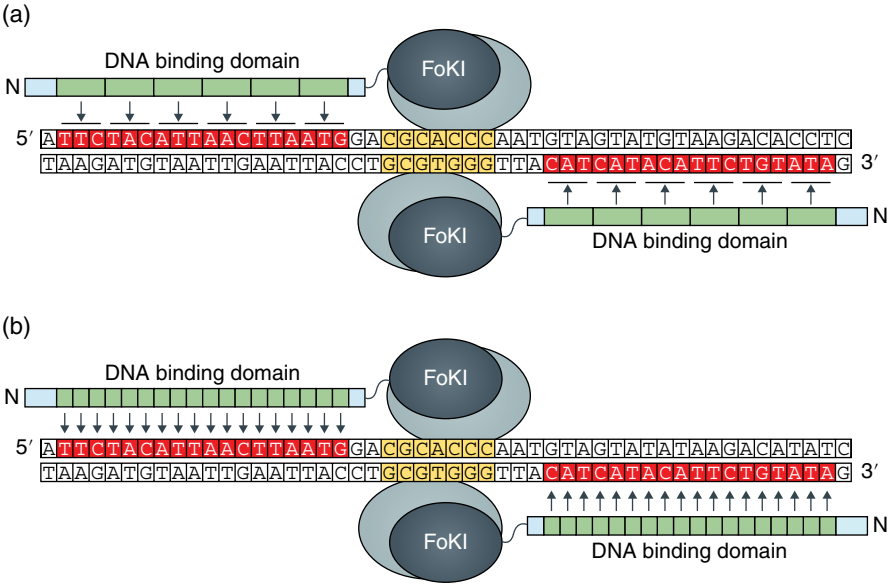


Figure 17.5. Engineered zinc finger nucleases (ZFNs; (a) and transcription activator-like effector nucleases (TALENs); (b) for targeted genome modification (reprinted with permission from Liu et al. 2013b). Each nuclease contains a custom-designed DNA-binding domain and the nonspecific DNA-cleavage domain of the *Fok*I endonuclease which has to dimerize for DNA cleavage within the spacer regions between the two binding sites. The spacer regions between the monomers of both nucleases are 5–7 bp and 6–40 bp in length, respectively.

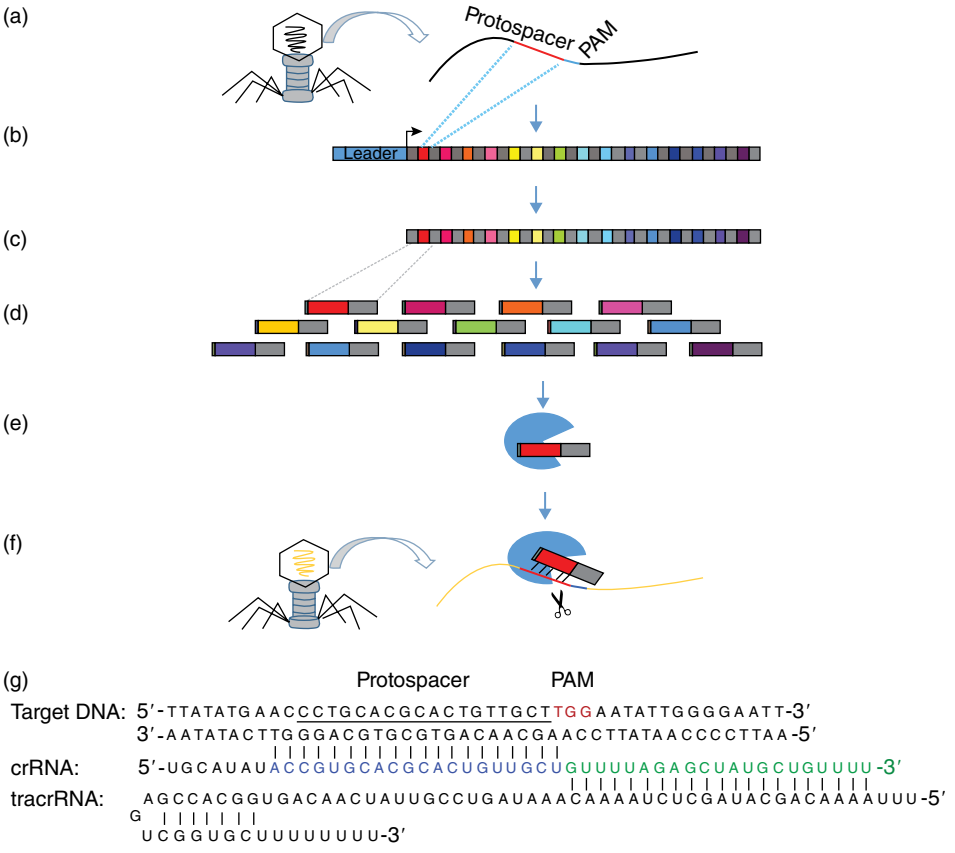


Figure 17.6. Outline of the CRISPR-Cas defense pathway (modified from Terns and Terns (2011) and Jinek et al. (2012)). (a) A short viral or plasmid DNA sequence (protospacer) upstream (for type II system) to the protospacer adjacent motif (PAM) is acquired and integrated into the host CRISPR locus adjacent to the leader sequence. (b) The CRISPR locus consists of invader-derived spacers with similar sizes (multiple colors) interspersed with short direct repeats (dark gray) and the leader sequence. (c) The transcription of the CRISPR locus using the leader as promoter produces pre-crRNA. (d) The pre-crRNA is processed to be mature crRNA, which typically contains an 8-necleotide repeat sequence at the 5′-end and a 20-nucleotide repeat sequence at the 3′-end. (e) Each crRNA binds to Cas9 protein (blue) with the ability to target different protospacers. (f) The binding of the crRNA-Cas9 ribonucleoprotein complex to a target sequence in the same or phylogentically closely related invader genome through base pairing leads to a double strand break in the target site. (g) Illustration of base pairing between crRNA and a target site on the foreign DNA, and between crRNA and tracrRNA in the type II CRISPR-Cas9 system.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.